



US009233894B2

(12) **United States Patent**  
**McPhee**(10) **Patent No.:** US 9,233,894 B2  
(45) **Date of Patent:** Jan. 12, 2016(54) **MICROBIAL DERIVED ISOPRENE AND METHODS FOR MAKING THE SAME**(71) Applicant: **AMYRIS, INC.**, Emeryville, CA (US)(72) Inventor: **Derek McPhee**, Emeryville, CA (US)(73) Assignee: **AMYRIS, INC.**, Emeryville, CA (US)

(\*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 89 days.

(21) Appl. No.: **13/887,381**(22) Filed: **May 6, 2013**(65) **Prior Publication Data**

US 2013/0221280 A1 Aug. 29, 2013

**Related U.S. Application Data**

(62) Division of application No. 13/629,623, filed on Sep. 28, 2012, now Pat. No. 8,492,605, which is a division of application No. 12/659,216, filed on Mar. 1, 2010, now Pat. No. 8,324,442.

(60) Provisional application No. 61/202,474, filed on Mar. 3, 2009.

(51) **Int. Cl.****C07C 11/16** (2006.01)  
**C07C 11/18** (2006.01)  
**C07C 7/09** (2006.01)  
**C12P 5/00** (2006.01)(52) **U.S. Cl.**CPC . **C07C 11/16** (2013.01); **C07C 7/09** (2013.01);  
**C12P 5/007** (2013.01); **Y02E 50/343** (2013.01)(58) **Field of Classification Search**

None

See application file for complete search history.

(56) **References Cited**

## U.S. PATENT DOCUMENTS

4,323,508 A \* 4/1982 Herrington et al. .... 549/505  
7,931,873 B2 4/2011 Lovegrove et al.  
2007/0272646 A1 \* 11/2007 Lancesseur et al. .... 215/224  
2010/0261942 A1 \* 10/2010 McPhee ..... 585/603

## FOREIGN PATENT DOCUMENTS

WO WO2008137092 11/2008  
WO WO2009100231 8/2009  
WO WO2010101855 9/2010

## OTHER PUBLICATIONS

Newman, et al., "High-level Production of Amorpha-4,11-Diene in a Two-Phase Partitioning Bioreactor of Metabolically Engineered *Escherichia coli*", Biotechnology and Bioengineering, vol. 95. No. 4, Nov. 5, 2006, pp. 684-691.

ISR

Written Opinion

IPRP

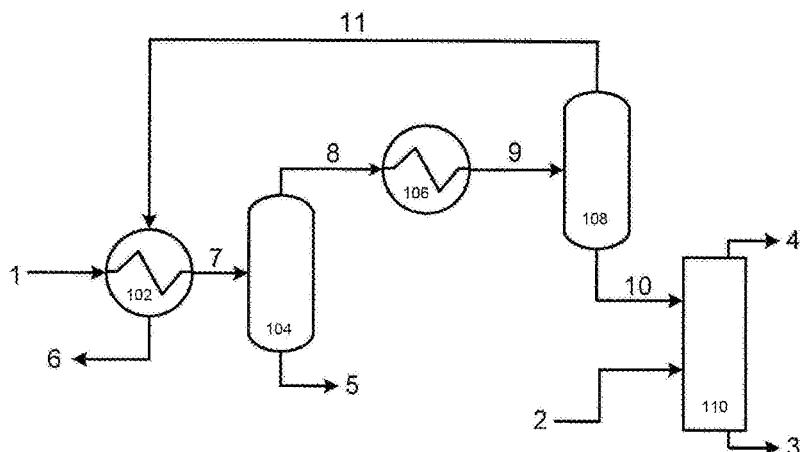
\* cited by examiner

Primary Examiner — Tam M Nguyen

(74) Attorney, Agent, or Firm — Kam. W. Law; Squire Patton Boggs (US) LLP

(57) **ABSTRACT**

Provided herein is a gaseous isoprene composition comprising isoprene, carbon dioxide and water, wherein the isoprene is in an amount between about 0.1% and about 15% by volume; wherein the carbon dioxide is in an amount between about 0.04% and about 35% by volume; wherein the water is in an amount greater than about 70% of its saturation amount. Also provided herein is a liquid isoprene composition comprising isoprene in an amount of at least 65% by weight and carbon dioxide in an amount between about 0.01% and about 1% by weight.

**16 Claims, 9 Drawing Sheets**

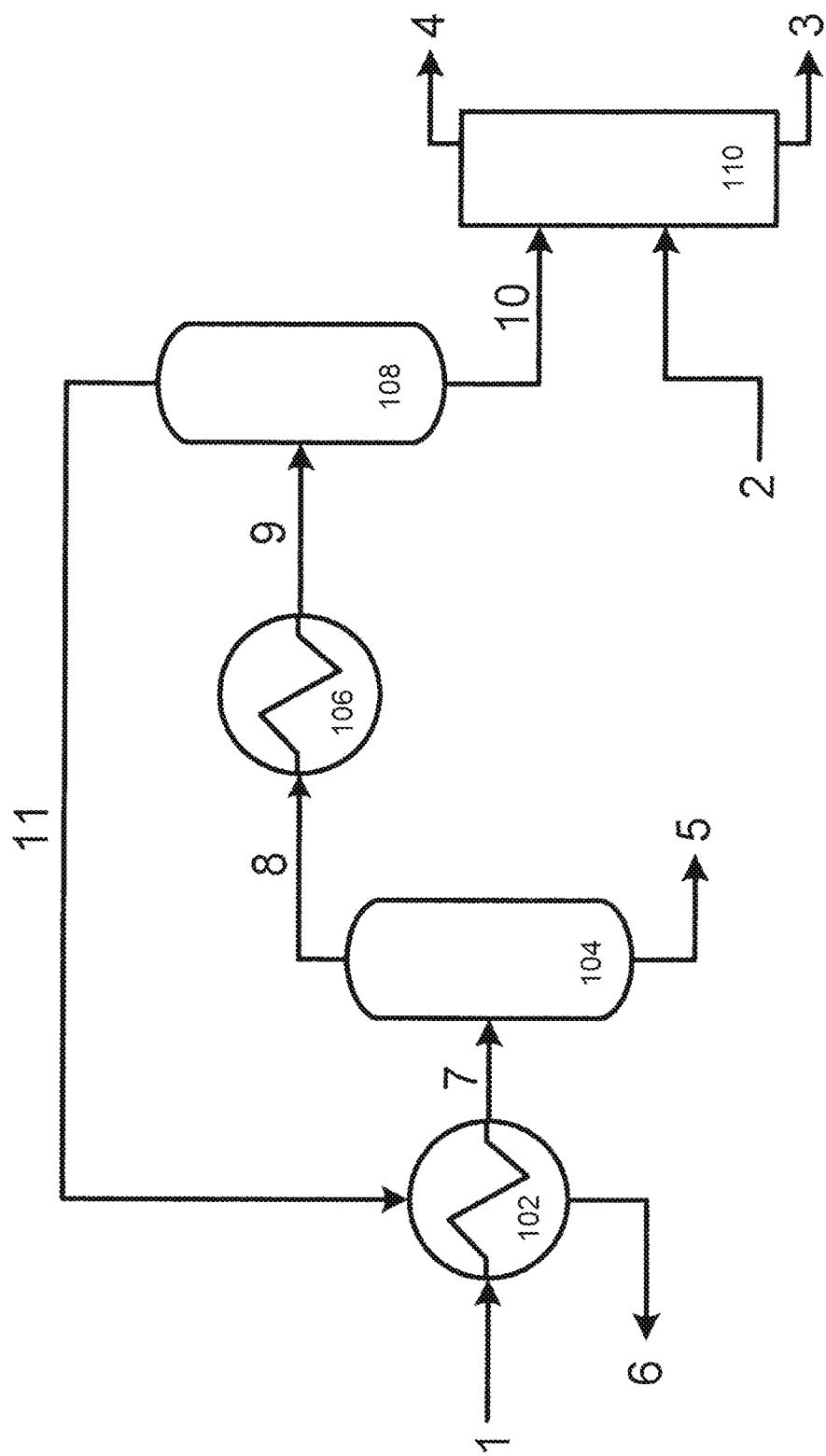


Figure 1

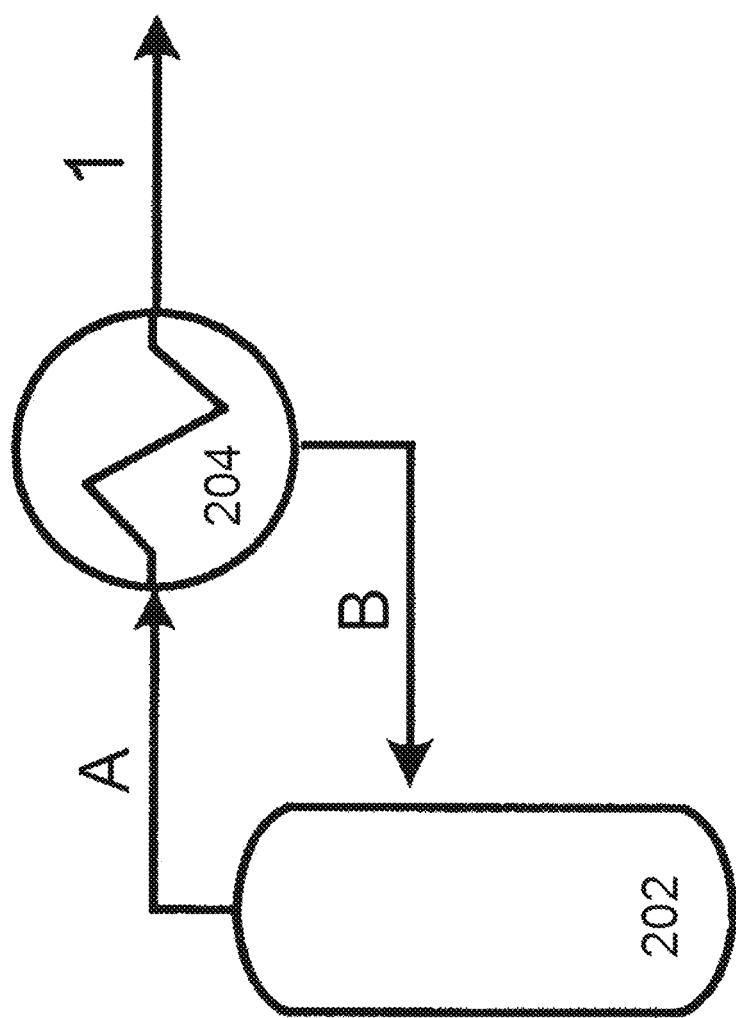
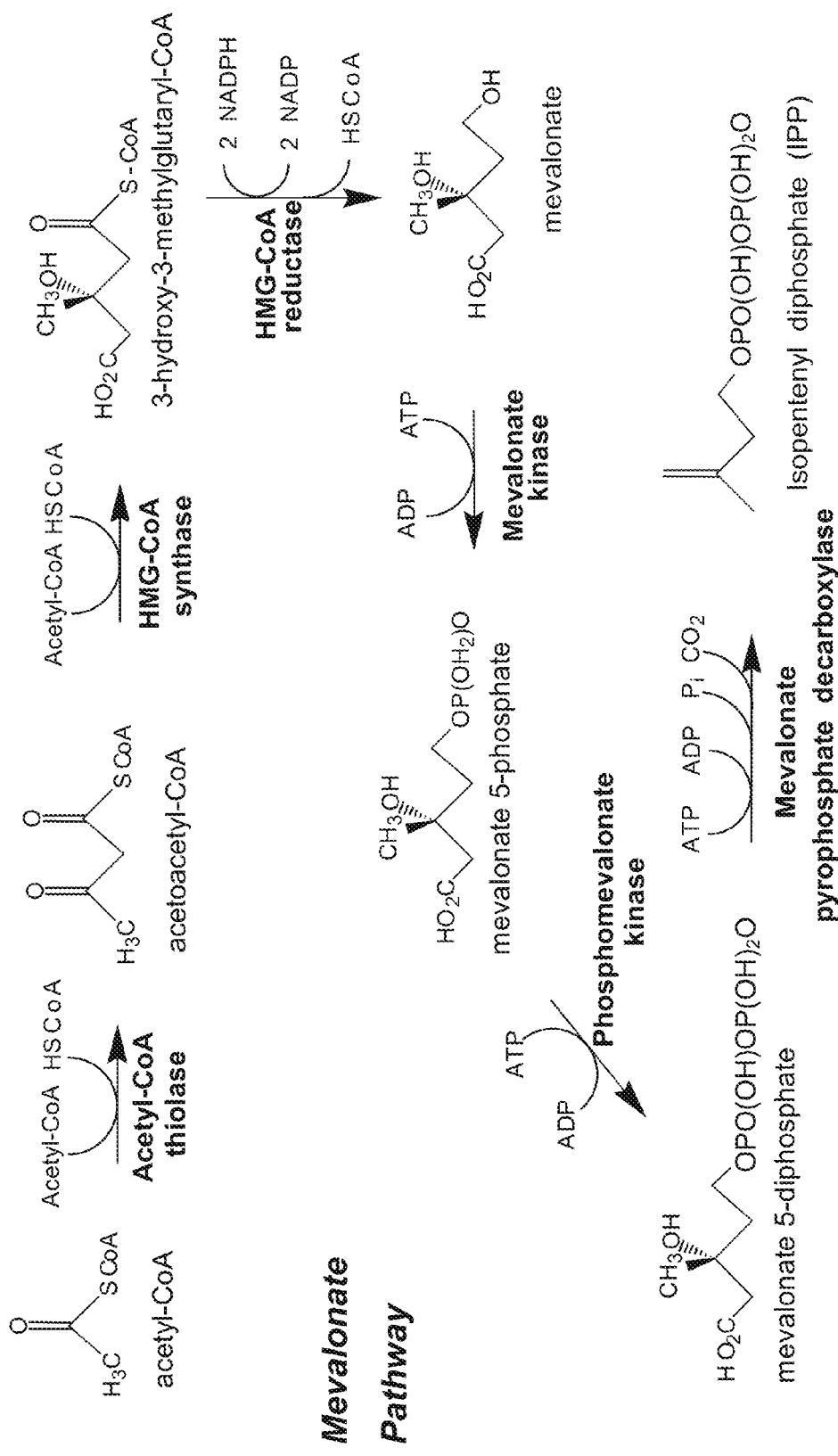
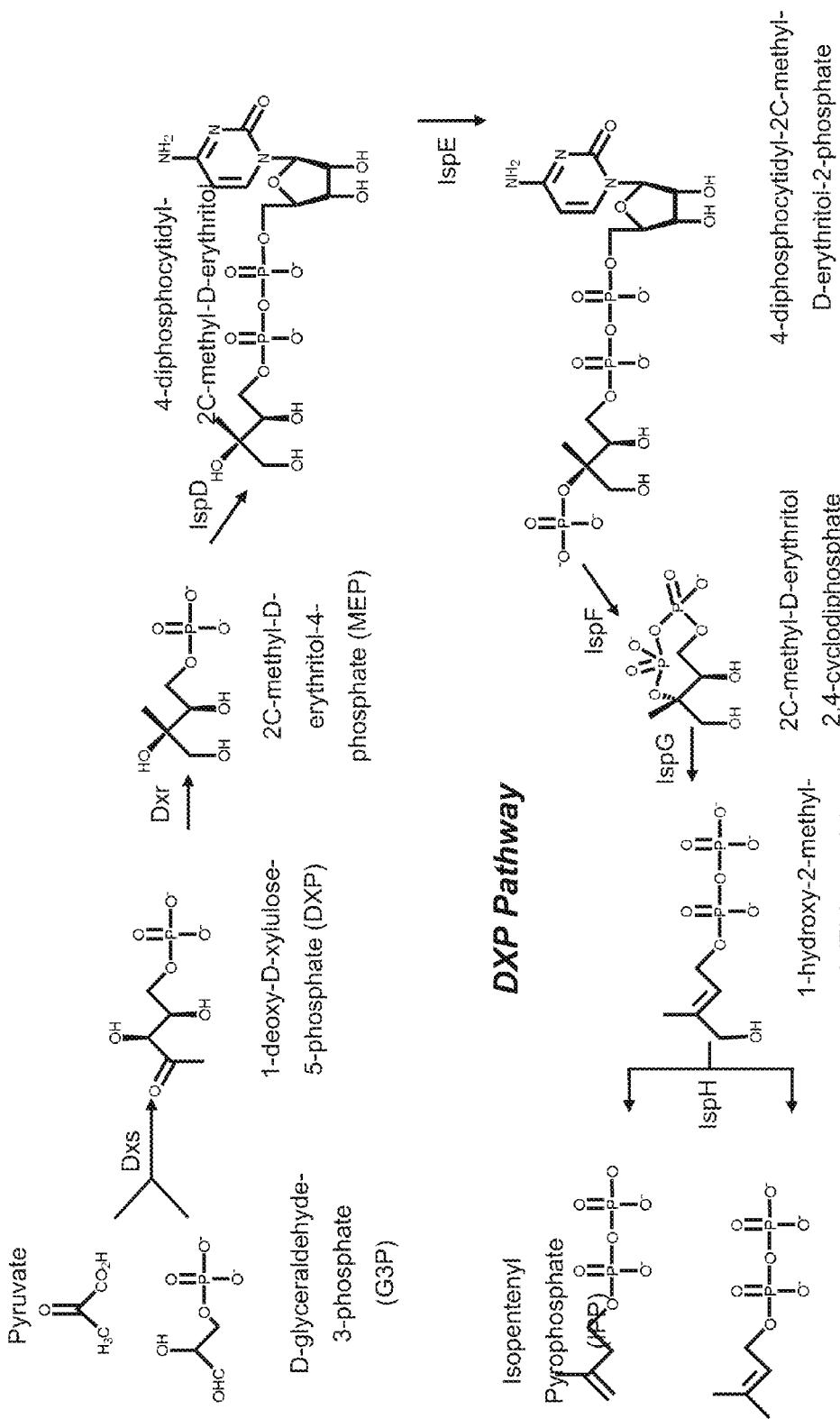
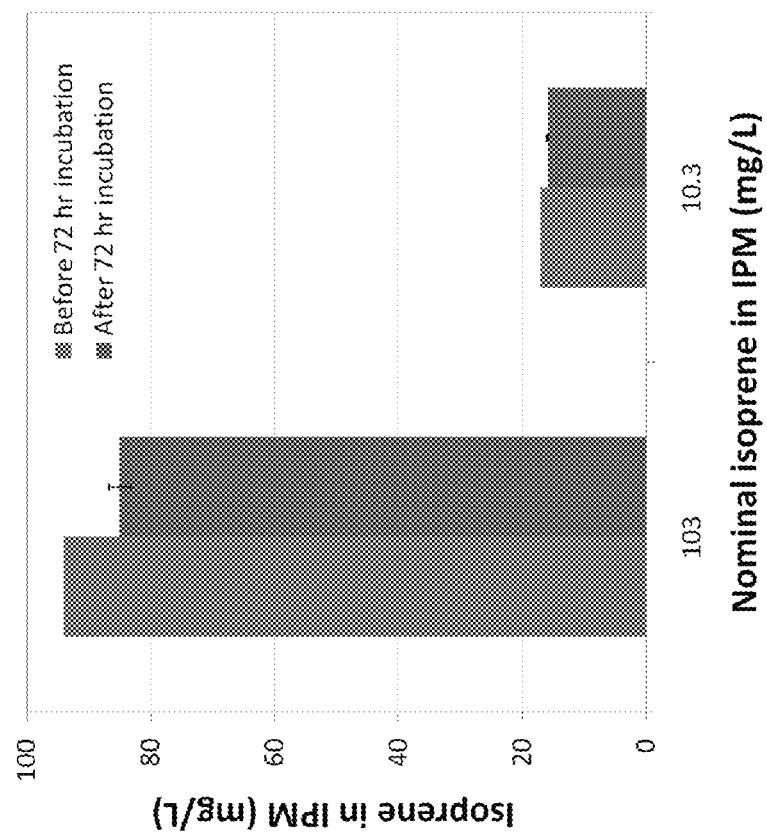


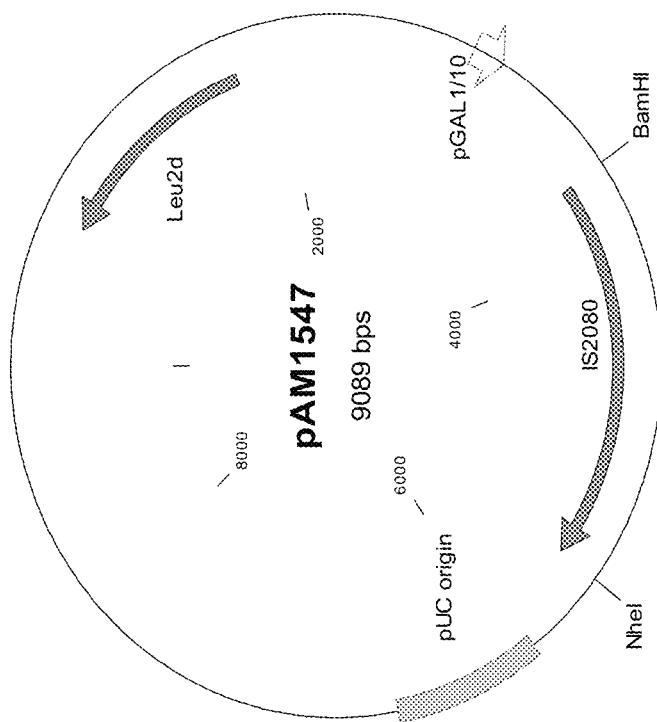
Figure 2

**Figure 3**

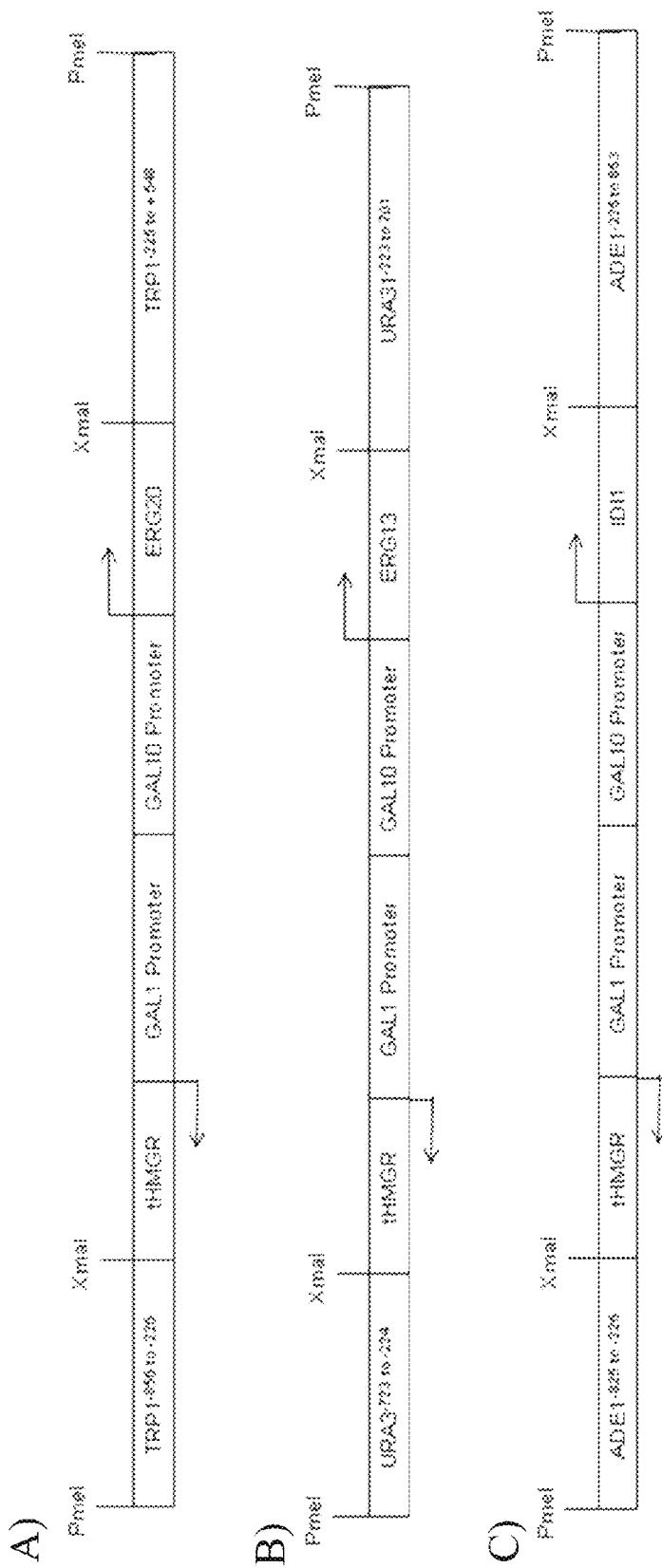
**Figure 4**



**Figure 5**



**Figure 6**

**Figure 7**

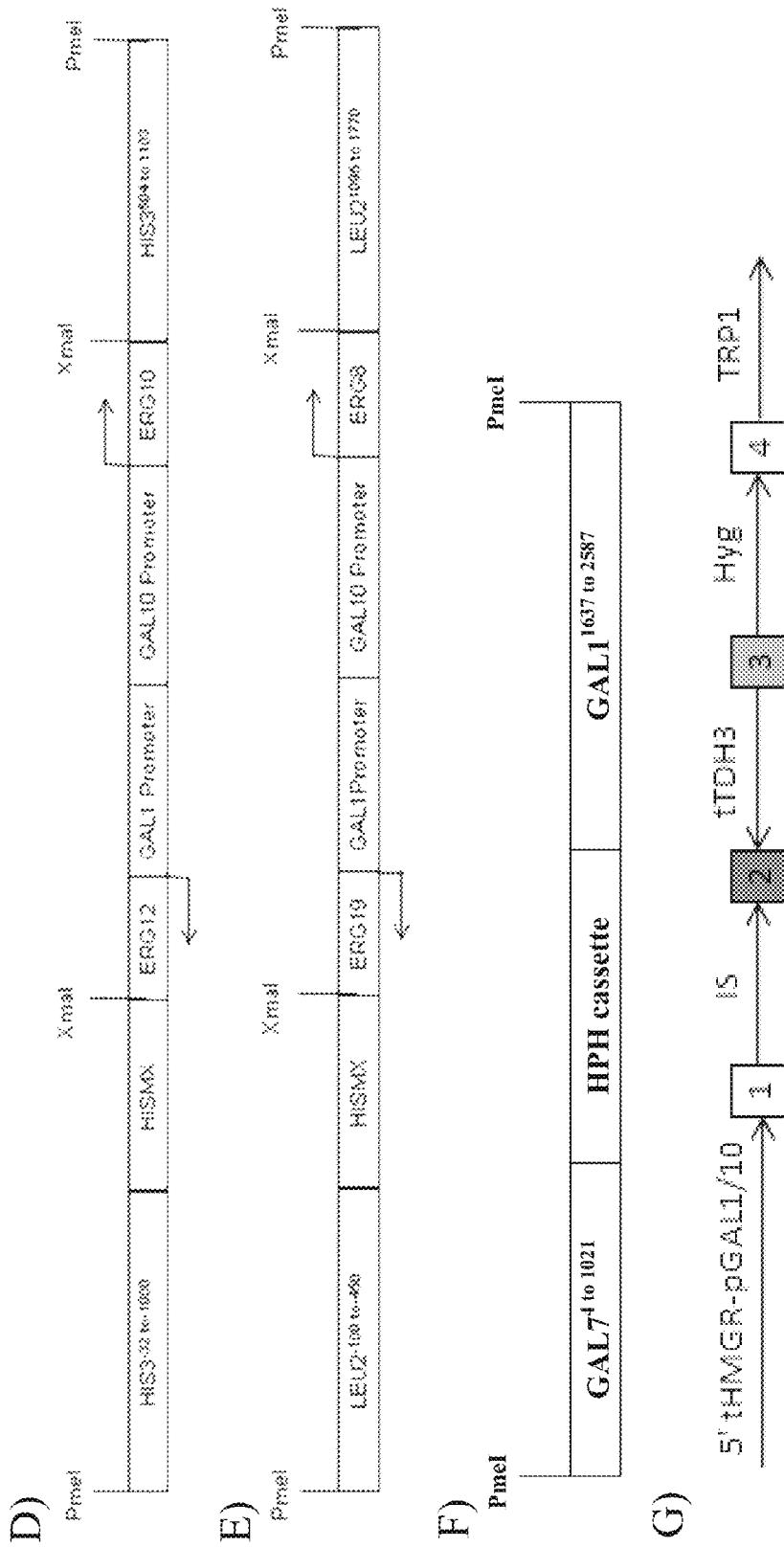
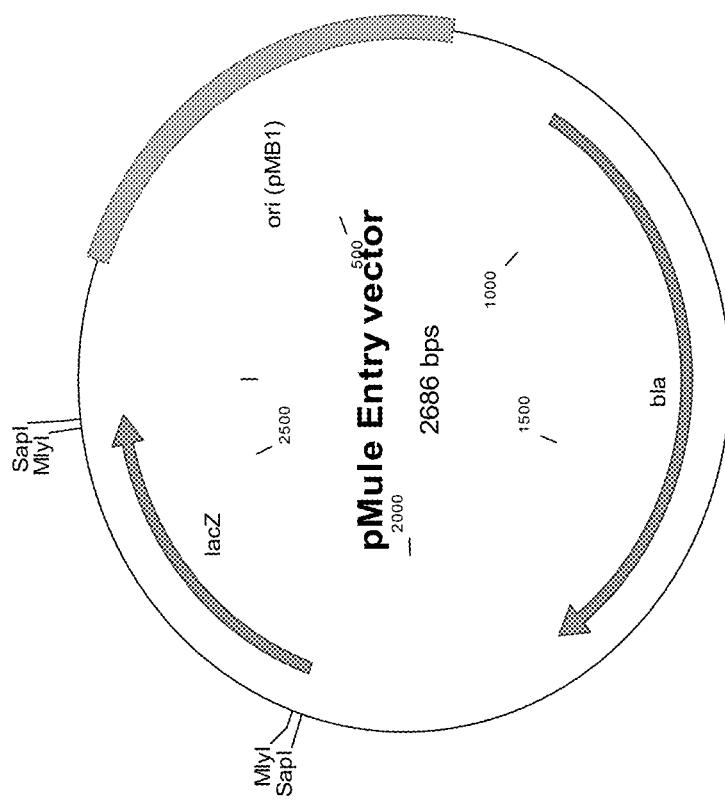


Figure 7 (continued)



**Figure 8**

**1****MICROBIAL DERIVED ISOPRENE AND METHODS FOR MAKING THE SAME****CROSS REFERENCE TO RELATED APPLICATION**

This is a divisional application of the U.S. Non-Provisional Application Ser. No. 13/629,623 now U.S. Pat. No. 8,492,605, which is a divisional application of the U.S. Non-Provisional application Ser. No. 12/659,216, now U.S. Pat. No. 8,324,442, which claims the benefit of priority from U.S. Provisional Application No. 61/202,474, filed Mar. 3, 2009, all of which in their entirety are incorporated herein by reference.

**BACKGROUND OF THE INVENTION**

Isoprene is a five carbon hydrocarbon (2-methyl-1,3-butadiene), that is an industrial chemical used in a range of industrial application such as tires, footwear, sporting goods, latex, tapes, labels, and medical disposables. Isoprene is also a natural compound produced in biological systems. While isoprene is made naturally in various organisms ranging from microbes to animals, most naturally occurring isoprene has traditionally been extracted from rubber plants. However, extraction yields are low and these quantities are far less than are required for many commercial applications. As a result, isoprene is primarily produced synthetically from petroleum sources, most often from ethylene using a steam cracking process.

Due to the growing concern for climate change and thus a need to make products we need more sustainably, there is an urgent need for bio- or renewable isoprene that will help meet global isoprene demands but that can be produced in a more environmentally friendly way. The current invention addresses this need.

**SUMMARY OF THE INVENTION**

The present invention provides microbial derived isoprene compositions and methods for making and purifying the same.

In one aspect of the invention, a gaseous isoprene composition is provided comprising: isoprene, carbon dioxide and water, wherein the isoprene is in an amount between about 0.1% and about 15% by volume; wherein the carbon dioxide is in an amount between about 0.04% and about 35% by volume; wherein the water is in an amount greater than about 70% of its saturation amount; and wherein the gaseous isoprene composition comprises 1 part per million or less than 1 part per million of any one of the following impurities: C<sub>2</sub>-C<sub>5</sub> alkynes, cyclopentadiene, piperylene, and 1,4-pentadiene.

In some embodiments, the gaseous composition comprises less than about 3% by weight of water. In certain embodiments, the gaseous composition further comprises oxygen in an amount between about 1% and about 20% by volume. In some embodiments, the gaseous composition further comprises nitrogen in an amount greater than about 50% by volume. In certain embodiments, the gaseous composition further comprises argon in an amount less than about 0.9% by volume or greater than about 1.0% by volume. In certain embodiments, the gaseous composition further comprises ethanol in an amount less than about 0.5% by volume or more than about 1% by volume. In certain embodiments, the gaseous composition further comprises 1 part per million or less of cyclopentadiene, piperylene, 1,4-pentadiene, or a combination thereof.

**2**

In another aspect a liquid isoprene composition is provided comprising: isoprene in an amount of at least 65% by weight and carbon dioxide in an amount between about 0.01% and about 1% by weight, wherein the isoprene composition comprises 1 part per million or less than 1 part per million of any one of the following impurities: C<sub>2</sub>-C<sub>5</sub> alkynes, cyclopentadiene, piperylene, and 1,4-pentadiene.

In some embodiments, the liquid isoprene composition disclosed herein comprises at least about 80% or at least about 95% isoprene by weight. In some embodiments, the liquid isoprene composition comprises less than about 1% by weight of water. In certain embodiments, the liquid isoprene composition further comprises nitrogen in an amount between about 0.001% and about 1% by weight. In some embodiments, the liquid isoprene composition further comprises ethanol in an amount greater than about 0.01% by weight. In certain embodiments, the carbon dioxide is in an amount between about 0.05% and about 1% by weight, or between about 0.1% and about 1% by weight, or between about 0.2% and about 0.7% by weight.

In another aspect, a method for making and purifying isoprene is provided. The method comprises:

- a. obtaining a first gaseous composition comprising isoprene and water wherein the gaseous composition comprises 1 part per million or less of C<sub>2</sub>-C<sub>5</sub> alkynes;
- b. flowing the first gaseous composition through a first chiller wherein the first chiller has a temperature of between about 10° C. and about -15° C. thereby resulting in a second gaseous composition and wherein the second gaseous composition comprises less water than the first gaseous composition;
- c. flowing the second gaseous composition through a second chiller wherein the second chiller has a temperature below -35° C.; and
- d. collecting the resulting liquid isoprene composition.

In another aspect, another method is provided. The method comprises:

- a. culturing a plurality of host cells capable of making isoprene;
- b. forming a first gaseous composition comprising isoprene and water wherein the water is present in an amount greater than about 70% of its saturation amount;
- c. subjecting the first gaseous composition to a first cooling step whereby substantially all of the water is removed from the first gaseous composition resulting a second gaseous composition; and
- d. subjecting the second gaseous composition to a second cooling step whereby a liquid isoprene composition is collected.

In another aspect, another method is provided. The method comprises:

- a. culturing a plurality of host cells capable of making isoprene;
- b. forming a first gaseous composition comprising isoprene and water wherein the water is present in an amount greater than about 70% of its saturation amount;
- c. subjecting the first gaseous composition to a first cooling step whereby substantially all of the water is removed from the first gaseous composition resulting a second gaseous composition;
- d. subjecting the second gaseous composition to a second cooling step whereby a liquid isoprene composition is collected; and
- e. optionally, contacting either the first gaseous composition, the second gaseous composition and/or the liquid

isoprene composition with a membrane containing modified zeolites or molecular sieve to provide a purified isoprene composition.

In another aspect, another method is provided. The method comprises:

- a. contacting a plurality of host cells capable of making isoprene in an aqueous medium wherein the aqueous medium is in contact with an immiscible organic liquid and the aqueous medium, the host cells, and the immiscible organic liquid is in a closed vessel; and
- b. culturing the host cells in the aqueous medium whereby the host cells make isoprene and the isoprene is captured in the immiscible organic liquid.

In another aspect, another method is provided. The method comprises:

- a. obtaining a first gaseous composition comprising:
  - i. isoprene in an amount between about 0.1% and about 15% by volume;
  - ii. carbon dioxide in an amount between about 0.04% and about 35% by volume;
  - iii. oxygen in an amount between about 1% and about 20% by volume;
  - iv. nitrogen in an amount greater than about 50% by volume;
  - v. argon in an amount less than about 0.9% by volume;
  - vi. water in an amount greater than about 70% of its saturation amount;
  - vii. 1 part per million or less of C2-C5 alkyne, cyclopentadiene, piperylene, and 1,4-pentadiene; and
  - viii. ethanol;

- b. flowing the first gaseous composition through a first chiller and an operably connected flash drum, wherein the first chiller has a temperature of between about 10° C. and about -15° C. thereby resulting in a second gaseous composition and wherein the second gaseous composition comprises less water than the first gaseous composition;
- c. flowing the second gaseous composition through a second chiller and an operably connected flash drum, wherein the second chiller has a temperature between about 35° C. and about -85° C.; and
- d. collecting the resulting liquid isoprene composition.

An isoprene production system comprising:

- a. a bioreactor capable of culturing a plurality of host cells;
- b. a first chiller and flash drum operably connected to the overhead stream of the bioreactor, the first chiller capable of operating in a temperature range of between 10° C. and -15° C.; and
- c. a second chiller and flash drum operably connected to the overhead stream exiting from the first chiller and flash drum, the second chiller capable of operating in a temperature below -35° C.

In yet another aspect, an isoprene production system is provided. The system comprises:

- a. a closed vessel;
- b. an aqueous medium, within the vessel, forming a first phase;
- c. a plurality of host cells, within the aqueous medium, capable of making isoprene; and,
- d. a liquid organic second phase, capable of capturing the isoprene made by the host cells, in contact with the first phase.

## BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a schematic representation of an exemplary separation system.

FIG. 2 is another schematic representation of another exemplary separation system.

FIG. 3 is a schematic representation of the mevalonate ("MEV") pathway for the production of isopentenyl diphosphate ("IPP").

FIG. 4 is a schematic representation of the DXP pathway for the production of IPP and dimethylallyl pyrophosphate ("DMAPP"). Dxs is 1-deoxy-D-xylulose-5-phosphate synthase; Dxr is 1-deoxy-D-xylulose-5-phosphate reductoisomerase (also known as IspC); IspD is 4-diphosphocytidyl-2C-methyl-D-erythritol synthase; IspE is 4-diphosphocytidyl-2C-methyl-D-erythritol synthase; IspF is 2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase; IspG is 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate synthase (IspG); and IspH is isopentenyl/dimethylallyl diphosphate synthase.

FIG. 5 shows the recovery of isoprene from an immiscible organic liquid (isopropyl myristate) from a closed fermentation system.

FIG. 6 shows a map of plasmid pAM1547.

FIGS. 7A-G show maps of the inserts of vectors pAM489, pAM491, pAM493, pAM495, pAM497, and pAM584, and of the integration cassette natA-P<sub>CTR3</sub><sup>-1 to -734</sup>.

FIG. 8 shows a map of the pMULE Entry vector.

## DETAILED DESCRIPTION OF THE INVENTION

### Definitions

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Reference is made here to a number of terms that shall be defined to have the following meanings:

"Bio-organic compound" refers to an organic compound having at least five carbon atoms that can be made by a host cell by taking a carbohydrate carbon source and converting the carbohydrate carbon source into the desired product.

"Deoxyxylulose 5-phosphate pathway" or "DXP pathway" is used herein to refer to the pathway that converts glyceraldehyde-3-phosphate and pyruvate to IPP and DMAPP. The DXP pathway is illustrated schematically in FIG. 4.

"Endogenous" refers to a substance or process that can occur naturally, e.g., in a non-recombinant host cell.

"Heterologous nucleic acid" as used herein refers to a nucleic acid wherein at least one of the following is true: (a) the nucleic acid is foreign ("exogenous") to (that is, not naturally found in) a given host cell; (b) the nucleic acid comprises

55 a nucleotide sequence that is naturally found in (that is, is "endogenous to") a given host cell, but the nucleotide sequence is produced in an unnatural (for example, greater than expected or greater than naturally found) amount in the cell; (c) the nucleic acid comprises a nucleotide sequence that differs in sequence from an endogenous nucleotide sequence, but the nucleotide sequence encodes the same protein (having the same or substantially the same amino acid sequence) and is produced in an unnatural (for example, greater than expected or greater than naturally found) amount in the cell; or (d) the nucleic acid comprises two or more nucleotide sequences that are not found in the same relationship to each other in nature (for example, the nucleic acid is recombinant).

60

## 5

"Host cell" and "microorganism" are used interchangeably herein to refer to any archae, bacterial, or eukaryotic living cell into which a heterologous nucleic acid can be or has been inserted. The term also relates to the progeny of the original cell, which may not necessarily be completely identical in morphology or in genomic or total DNA complement to the original parent, due to natural, accidental, or deliberate mutation.

"Isoprenoid" and "isoprenoid compound" are used interchangeably herein and refer to a compound derivable from isopentenyl diphosphate.

"Isolate" and "isolating" when referred to a bio-organic compound is the enrichment of the amount of the bio-organic compound in a composition. Consequently, the amount of the bio-organic compound in a composition after the bio-organic compound has been isolated or subject to an isolating step is greater than the amount present in the composition prior to such step.

"Mevalonate pathway" or "MEV pathway" is used herein to refer to the biosynthetic pathway that converts acetyl-CoA to IPP. The MEV pathway is illustrated schematically in FIG. 3.

"Naturally occurring" as applied to a nucleic acid, an enzyme, a cell, or an organism, refers to a nucleic acid, enzyme, cell, or organism that is found in nature. For example, a polypeptide or polynucleotide sequence that is present in an organism that can be isolated from a source in nature and that has not been intentionally modified by a human in the laboratory is naturally occurring.

"Optional" or "optionally" means that the subsequently described feature or structure may or may not be present, or that the subsequently described event or circumstance may or may not occur, and that the description includes instances where a particular feature or structure is present and instances where the feature or structure is absent, or instances where the event or circumstance occurs and instances where the event or circumstance does not occur.

"Pyrophosphate" is used interchangeably herein with "diphosphate".

As used herein, a composition that is a "substantially pure" compound is substantially free of one or more other compounds, i.e., the composition contains greater than 80 vol. %, greater than 90 vol. %, greater than 95 vol. %, greater than 96 vol. %, greater than 97 vol. %, greater than 98 vol. %, greater than 99 vol. %, greater than 99.5 vol. %, greater than 99.6 vol. %, greater than 99.7 vol. %, greater than 99.8 vol. %, greater than 99.9 vol. % of the compound; or less than 20 vol. %, less than 10 vol. %, less than 5 vol. %, less than 4 vol. %, less than 3 vol. %, less than 2 vol. %, less than 1 vol. %, less than 0.5 vol. %, less than 0.1 vol. %, or less than 0.01 vol. % of the one or more other compounds, based on the total volume of the composition.

In the following description, all numbers disclosed herein are approximate values, regardless whether the word "about" or "approximate" is used in connection therewith. They may vary by 1 percent, 2 percent, 5 percent, or, sometimes, 10 to 20 percent. Whenever a numerical range with a lower limit, RL and an upper limit, RU, is disclosed, any number falling within the range is specifically disclosed. In particular, the following numbers within the range are specifically disclosed:  $R = RL + k^*(RU - RL)$ , wherein k is a variable ranging from 1 percent to 100 percent with a 1 percent increment, i.e., k is 1 percent, 2 percent, 3 percent, 4 percent, 5 percent, . . . , 50 percent, 51 percent, 52 percent, . . . , 95 percent, 96 percent, 97 percent, 98 percent, 99 percent, or 100 percent. Moreover, any numerical range defined by two R numbers as defined in the above is also specifically disclosed.

## 6

In addition to the definitions above, certain compounds described herein have one or more double bonds that can exist as either the Z or E isomer. The invention in certain embodiments encompasses these compounds as individual isomers in a substantially pure form as well as mixtures of various isomers, e.g., racemic mixtures of stereoisomer.

## Current Sources of Isoprene

Isoprene currently is made naturally by rubber plants (typically *Hevea brasiliensis*) or is made synthetically from petroleum sources. When made naturally, the sap-like extract (known as latex and is a polymerized version of isoprene) is collected from the rubber plants and is the primary source of natural rubber. Because latex and natural rubber can be of varying quality (irregular molecular distribution), the synthetic analog of natural rubber or polyisoprene is often preferred due to its higher uniformity.

Chemically synthesized isoprene is made primarily from petroleum sources. The most common method involves stream cracking a petroleum stream to make ethylene which in turn is subsequently converted into isoprene. Other methods for making isoprene include isobutylene carbonylation and isopentane dehydrogenation. The resulting isoprene is produced and sold in different concentrations. Crude isoprene has a purity between 15% and 65%. Refined isoprene is defined as isoprene having a purity between 65% and 95%. High purity isoprene is defined as isoprene having a purity between 95% and 99.5%. Polymer grade isoprene is isoprene with a purity exceeding 99.5%.

As a consequence of how it is made, synthetic isoprene contains a number of impurities including various acetylenes and dienes such as cyclopentadiene and piperylene. Although these catalysts are undesirable as they inhibit polymerization, it is not often economical to entirely eliminate them and the purity of the isoprene is matched to the desired end product. For example, the isoprene purity required to make butyl rubber is substantially less than required to make SIS polymers (polymer grade).

## Microbially Derived Gaseous Isoprene Compositions

The present invention provides microbial derived isoprene compositions and methods for making and purifying the same. Microbial-derived isoprene compositions differ from petroleum derived sources in that the compositions include virtually none of the following impurities: C<sub>2</sub>-C<sub>5</sub> alkynes; cyclopentadiene, piperylene and 1,4-pentadiene.

In one aspect of the invention, a gaseous isoprene composition is provided. The composition comprises isoprene and water wherein the water is present in an amount that is at least about 70% of its saturation amount and the composition comprises 1 part per million or less of C<sub>2</sub>-C<sub>5</sub> alkynes. Illustrative examples of C<sub>2</sub>-C<sub>5</sub> alkynes include acetylene, isopropylacetylene, 1-pentyne, 2-pentyne, and 2-butyne.

In another aspect of the invention, another gaseous isoprene composition is provided. The composition comprises isoprene and water wherein the water is present in an amount that is at least about 70% of its saturation amount and the composition comprises 1 part per million or less of cyclopentadiene.

In another aspect of the invention, another gaseous isoprene composition is provided. The composition comprises isoprene and water wherein the water is present in an amount that is at least about 70% of its saturation amount and the composition comprises 1 part per million or less of piperylene.

In another aspect of the invention, another gaseous isoprene composition is provided. The composition comprises isoprene and water wherein the water is present in an amount

that is at least about 70% of its saturation amount and the composition comprises 1 part per million or less of 1,4-pentadiene.

In another aspect of the invention, another gaseous isoprene composition is provided. The composition comprises isoprene and water wherein the water is present in an amount that is at least about 70% of its saturation amount and the composition comprises 1 part per million or less of each of C<sub>2</sub>-C<sub>5</sub> alkynes, cyclopentadiene, piperylene, and 1,4-pentadiene.

In some embodiments, the gaseous isoprene composition comprises isoprene that is present between about 0.1% and about 15% by volume. In other embodiments, the isoprene is present between about 1 and 10% by volume. In still other embodiments, the isoprene is present between about 1 and 5% by volume. In yet other embodiments, the isoprene is present between about 5% and about 10% by volume. In further embodiments, the isoprene is present between in an amount greater than about 10% by volume.

In other embodiments, the gaseous isoprene composition comprises water in an amount that is greater than about 70%, 75%, 80%, 85%, 90%, 95% and 99% of its saturation amount. In still other embodiments, the gaseous isoprene composition comprises saturated water.

In other embodiments, the gaseous isoprene composition further comprises carbon dioxide that is present in an amount that is greater than about 0.04% by volume. In still other embodiments, the carbon dioxide is present in an amount that is greater than about 0.05%, 0.1%, 0.5%, 1.0%, and 5% by volume. In further embodiments, the carbon dioxide is present in an amount that is greater than about 10%, about 20%, about 30% by volume. In still further embodiments, the carbon dioxide is present in an amount that is between about 1% and about 35% by volume. In still other embodiments, the carbon dioxide is present in an amount that is between about 10% and about 30% by volume.

In other embodiments, the gaseous isoprene composition further comprises oxygen. In some embodiments, the oxygen is present in an amount that is less than about 20.9% by volume. In other embodiments, the oxygen is present in an amount that is between about 1% by volume and about 20% by volume. In other embodiments, the oxygen is present in an amount that is between about 8% and about 15% by volume. In other embodiments, the oxygen is present in an amount that is less than about 20%, 15%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, and 2%. In still other embodiments, the oxygen is present in an amount that is less than about 1% by volume. In further embodiments, the oxygen is present between about 1% and about 15% by volume. In still further embodiments, the oxygen is present between about 5% and about 15% by volume.

In other embodiments, the gaseous isoprene composition further comprises nitrogen. In some embodiments, the nitrogen is present in an amount between about 50% and about 75% by volume. In further embodiments, the nitrogen is present in an amount that is greater than about 70%. In other embodiments, the nitrogen is present in an amount that is greater than about 75%, 76%, 77%, 78%, 79%, and 80%.

In other embodiments, the gaseous isoprene composition further comprises argon. In some embodiments, the argon is present in an amount that is less than about 0.9% by volume. In other embodiments, the argon is present in an amount that is greater than about 1.0% by volume.

In other embodiments, the gaseous isoprene composition further comprises ethanol. In some embodiments, the ethanol is present in an amount that is less than about 0.5% by volume. In other embodiments, the ethanol is present in an amount that is more than about 1% by volume.

In other embodiments, the microbial-derived gaseous isoprene composition may comprise: isoprene in an amount between about 0.1% and about 15% by volume; water in an amount that is greater than about 70% of its saturation amount; carbon dioxide in an amount that is between about 0.04% and about 35% by volume; oxygen in an amount that is between about 1% and about 20% by volume; nitrogen in an amount that is greater than about 50% by volume; argon in an amount that is less than about 0.9% by volume; ethanol in an amount that is less than about 0.5% by volume; 1 part per million or less of C<sub>2</sub>-C<sub>5</sub> alkynes; 1 part per million or less of cyclopentadiene; 1 part per million or less of piperylene; and 1 part per million or less of 1,4-pentadiene.

In other embodiments, the microbial-derived gaseous isoprene composition may comprise: isoprene in an amount between about 0.1% and about 15% by volume; water in an amount that is greater than about 70% of its saturation amount; carbon dioxide in an amount that is between about 0.04% and about 35% by volume; oxygen in an amount that is between about 1% and about 20% by volume; nitrogen in an amount that is greater than about 50% by volume; argon in an amount that is greater than about 1.0% by volume; ethanol in an amount that is more than about 1% by volume; 1 part per million or less of C<sub>2</sub>-C<sub>5</sub> alkynes; 1 part per million or less of cyclopentadiene; 1 part per million or less of piperylene; and 1 part per million or less of 1,4-pentadiene.

In certain other embodiments, another gaseous isoprene composition is provided. This composition comprises:

- a. isoprene in an amount between about 0.1% and 15% by volume;
- b. carbon dioxide in an amount between about 1% and 35% by volume; and,

c. water in an amount that is greater than about 70% of its saturation amount and wherein the gaseous isoprene composition comprises 1 part per million or less of C<sub>2</sub>-C<sub>5</sub> alkynes. In other embodiments, the gaseous isoprene composition comprises 1 part per million or less of each of C<sub>2</sub>-C<sub>5</sub> alkynes, cyclopentadiene, piperylene, and 1,4-pentadiene. In still other embodiments, the gaseous isoprene composition comprises saturated water. In yet other embodiments, the gaseous isoprene composition further comprises oxygen in an amount between about 8% and about 15% by volume or nitrogen in an amount between about 50% and 75% by volume or both.

The temperature of the above described gaseous compositions is at least 30° C. In some cases, the temperature is between about 30° C. and about 60° C. In other cases, the temperature is between about 30° C. and about 38° C.

The pressure of the above described gaseous compositions is between about 1 and about 2.5 atmospheres.

For some of the above described gaseous compositions, the temperature is between about 30° C. and about 35° C. and is at a pressure between about 1 and about 2.5 atmospheres.

#### Microbially Derived Liquid Isoprene Compositions

Using the methods described herein, the gaseous isoprene compositions of the present invention can be further purified to liquid isoprene. Thus in another aspect of the invention, a liquid isoprene composition is provided that results from the inventive methods. The resulting liquid isoprene composition comprises at least 65% isoprene by weight and wherein the liquid isoprene composition comprises 1 part per million or less of C<sub>2</sub>-C<sub>5</sub> alkynes, cyclopentadiene, piperylene, and 1,4-pentadiene.

In some embodiments, the liquid isoprene composition comprises at least about 70%, 75%, 80%, 85%, and 90% isoprene by weight. In other embodiments, the liquid isoprene comprises at least about 95%, 96%, 97%, 98%, 99% and 99.5% isoprene by weight. In still other embodiments, the

liquid isoprene composition comprises isoprene in an amount that is greater than about 99.5% by weight.

In other embodiments, the liquid isoprene composition further comprises carbon dioxide. In some embodiments, the carbon dioxide is present in an amount that is between about 0.01% by weight and about 1% by weight. In other embodiments, the carbon dioxide is present in an amount that is between about 0.05% and about 1% by weight. In further embodiments, the carbon dioxide is present in an amount that is between about 0.1% and about 1% by weight. In still further embodiments, the carbon dioxide is present in an amount that is between about 0.2% and about 0.7% by weight.

In other embodiments, the liquid isoprene composition further comprises nitrogen. In some embodiments, the nitrogen is present in an amount that is between about 0.001% by weight and about 1% by weight. In other embodiments, the carbon dioxide is present in an amount that is between about 0.01% and about 0.5% by weight. In further embodiments, the carbon dioxide is present in an amount that is between about 0.05% and about 0.5% by weight.

In other embodiments, the liquid isoprene composition further comprises ethanol. In some embodiments, the ethanol is present in an amount that is greater than about 0.01% by weight. In other embodiments, the ethanol is present in an amount that is greater than about 0.1%, 0.2%, 0.3%, 0.4%, 0.5%, 0.6%, 0.7%, 0.8%, and 0.9% by weight. In further embodiments, the ethanol is present in an amount that is greater than about 1% by weight.

In other embodiments, the liquid isoprene composition may comprise water in an amount that is less than about 1%, 0.5%, 0.1%, and 0.05% by weight. In other embodiments, the liquid isoprene composition may comprise water in an amount that is less than about 500 ppm, 250 ppm, 100 ppm, and 50 ppm by weight. In other embodiments, the liquid isoprene composition may comprise water in an amount, by weight, that is less than the level of detection.

In other embodiments, the microbial-derived liquid isoprene composition may comprise: isoprene in an amount of at least about 65% to an amount greater than about 99.5% by weight; carbon dioxide in an amount that is between about 0.01% and about 1% by weight; nitrogen in an amount that is between about 0.001% and about 1% by weight; ethanol in an amount greater than about 0.01% to an amount greater than about 1% by weight; water in an amount that is less than about 1% by weight to an amount that is less than the level of detection; C<sub>2</sub>-C<sub>5</sub> alkynes in an amount 1 part per million or less; cyclopentadiene in an amount 1 part per million or less; piperylene in an amount 1 part per million or less; and 1,4-pentadiene in an amount 1 part per million or less.

In certain other embodiments, another liquid isoprene composition is provided. This composition comprises:

- a. isoprene in an amount greater than about 65% by weight;
- b. ethanol in an amount greater than about 0.01% by weight; and,
- c. carbon dioxide in an amount between about 0.01% and about 1% by weight wherein the liquid isoprene composition comprises 1 part per million or less of C<sub>2</sub>-C<sub>5</sub> alkynes, cyclopentadiene and piperylene C<sub>2</sub>-C<sub>5</sub> alkynes. In some embodiments, the isoprene is present in an amount greater than about 85% by weight. In still other embodiments, the isoprene is present in an amount greater than about 90% by weight. In further embodiments, the isoprene is present in an amount greater than about 90% by weight and ethanol is present in an amount that is between about 0.01% and about 1% by weight.

For some of the above described liquid isoprene compositions, the compositions have a temperature below -35° C. and a pressure between 0.01 and about 2 atmospheres. In other

embodiments, the compositions have a temperature between -45° C. and about -85° C. and a pressure below about 1 atmosphere. In still further embodiments, the compositions have a temperature below -45° C. and a pressure below about 0.5 atmosphere.

#### Microbial Host Cells

Any microbial host cells capable of making isoprene can be used in the methods herein which would result in the inventive isoprene compositions.

Illustrative examples of suitable host cells are microbes that have been shown to make isoprene naturally. These strains include those described by U.S. Pat. No. 5,849,970 and include: *Bacillus amyloliquefaciens*; *Bacillus cereus*; *Bacillus subtilis* 6051; *Basillus substillis* 23059; *Bacillus substillis* 23856; *Micrococcus luteus*; *Rhococcus rhodochrous*; *Acinetobacter calcoaceticus*; *Agrobacterium rhizogenes*; *Escherichia coli*; *Erwinia herbicola*; *Pseudomonas aeruginosa*; and *Psuedomonas citronellolis*. However, microbes that make isoprene naturally are produced at extremely low levels.

Isoprene is made from isopentenyl pyrophosphate (IPP) by isoprene synthase. Because all microbial host cells are capable of making IPP, any host cells can be made to make isoprene by the insertion of isoprene synthase into its genome. Illustrative examples of suitable nucleotide sequences include but are not limited to: (EF638224, *Populus alba*); (AJ294819, *Populus alba* × *Populus tremula*); (AM410988, *Populus nigra*); (AY341431, *Populus tremuloides*); (EF147555, *Populus trichocarpa*); and (AY316691, *Pueraria montana* var. *lobata*). The addition of a heterologous isoprene synthase to a microbial host cells that make isoprene naturally will improve isoprene yields of natural isoprene producers as well.

Any suitable microbial host cell can be genetically modified to make isoprene. A genetically modified host cell is one in which nucleic acid molecules have been inserted, deleted or modified (i.e., mutated; e.g., by insertion, deletion, substitution, and/or inversion of nucleotides), to produce isoprene. Illustrative examples of suitable host cells include any archae, bacterial, or eukaryotic cell. Examples of archae cells include, but are not limited to those belonging to the genera: *Aeropyrum*, *Archaeoglobus*, *Halobacterium*, *Methanococcus*, *Methanobacterium*, *Pyrococcus*, *Sulfolobus*, and *Thermoplasma*. Illustrative examples of archae species include but are not limited to: *Aeropyrum pernix*, *Archaeoglobus fulgidus*, *Methanococcus jannaschii*, *Methanobacterium thermoadutrophicum*, *Pyrococcus abyssi*, *Pyrococcus horikoshii*, *Thermoplasma acidophilum*, *Thermoplasma volcanium*.

Examples of bacterial cells include, but are not limited to those belonging to the genera: *Agrobacterium*, *Alicyclobacillus*, *Anabaena*, *Anacystis*, *Arthrobacter*, *Azobacter*, *Bacillus*, *Brevibacterium*, *Chromatium*, *Clostridium*, *Corynebacterium*, *Enterobacter*, *Erwinia*, *Escherichia*, *Lactobacillus*, *Lactococcus*, *Mesorhizobium*, *Methylobacterium*, *Microbacterium*, *Phormidium*, *Pseudomonas*, *Rhodobacter*, *Rhodopseudomonas*, *Rhodospirillum*, *Rhodococcus*, *Salmonella*, *Scenedesmus*, *Serratia*, *Shigella*, *Staphylococcus*, *Streptomyces*, *Synneccoccus*, and *Zymomonas*.

Illustrative examples of bacterial species include but are not limited to: *Bacillus subtilis*, *Bacillus amyloliquefacines*, *Brevibacterium ammoniagenes*, *Brevibacterium immarophilum*, *Clostridium beijerinckii*, *Enterobacter sakazakii*, *Escherichia coli*, *Lactococcus lactis*, *Mesorhizobium loti*, *Pseudomonas aeruginosa*, *Pseudomonas mevalonii*, *Pseudomonas pudica*, *Rhodobacter capsulatus*, *Rhodobacter sphaeroides*, *Rhodospirillum rubrum*, *Salmonella enterica*,

*Salmonella typhi*, *Salmonella typhimurium*, *Shigella dysenteriae*, *Shigella flexneri*, *Shigella sonnei*, *Staphylococcus aureus*, and the like.

In general, if a bacterial host cell is used, a non-pathogenic strain is preferred. Illustrative examples of species with non-pathogenic strains include but are not limited to: *Bacillus subtilis*, *Escherichia coli*, *Lactibacillus acidophilus*, *Lactobacillus helveticus*, *Pseudomonas aeruginosa*, *Pseudomonas mevalonii*, *Pseudomonas putida*, *Rhodobacter sphaeroides*, *Rodobacter capsulatus*, *Rhodospirillum rubrum*, and the like.

Examples of eukaryotic cells include but are not limited to fungal cells. Examples of fungal cells include, but are not limited to those belonging to the genera: *Aspergillus*, *Candida*, *Chrysosporium*, *Cryptococcus*, *Fusarium*, *Kluyveromyces*, *Neotyphodium*, *Neurospora*, *Penicillium*, *Pichia*, *Saccharomyces*, *Trichoderma* and *Xanthophyllomyces* (formerly *Phaffia*).

Illustrative examples of eukaryotic species include but are not limited to: *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Candida albicans*, *Chrysosporium lукновense*, *Fusarium graminearum*, *Fusarium venenatum*, *Kluyveromyces lactis*, *Neurospora crassa*, *Pichia angusta*, *Pichia finlandica*, *Pichia kodamae*, *Pichia membranaefaciens*, *Pichia methanolicus*, *Pichia opuntiae*, *Pichia pastoris*, *Pichia piperi*, *Pichia quercuum*, *Pichia salicaria*, *Pichia thermotolerans*, *Pichia trehalophila*, *Pichia stipitis*, *Streptomyces ambofaciens*, *Streptomyces aureofaciens*, *Streptomyces aureus*, *Saccaromyces bayanus*, *Saccaromyces boulardii*, *Saccharomyces cerevisiae*, *Streptomyces fungicidicus*, *Streptomyces griseochromogenes*, *Streptomyces griseus*, *Streptomyces lividans*, *Streptomyces olivogriseus*, *Streptomyces rameus*, *Streptomyces tanashiensis*, *Streptomyces vinaceus*, *Trichoderma reesei* and *Xanthophyllomyces dendrorhous* (formerly *Phaffia rhodozyma*).

In general, if a eukaryotic cell is used, a non-pathogenic strain is preferred. Illustrative examples of species with non-pathogenic strains include but are not limited to: *Fusarium graminearum*, *Fusarium venenatum*, *Pichia pastoris*, *Saccharomyces boulardi*, and *Saccharomyces cerevisiae*.

In some embodiments, the host cells of the present invention have been designated by the Food and Drug Administration as GRAS or Generally Regarded As Safe. Illustrative examples of such strains include: *Bacillus subtilis*, *Lactibacillus acidophilus*, *Lactobacillus helveticus*, and *Saccharomyces cerevisiae*.

In addition to the heterologous nucleic acid encoding an isoprene synthase, the microbial host cell can be further modified to increase isoprene yields. These modifications include but are not limited to the expression of one or more heterologous nucleic acid molecules encoding one or more enzymes in the mevalonate or DXP pathways.

#### MEV Pathway

A schematic representation of the MEV pathway is described in FIG. 3. In general, the pathway comprises six steps.

In the first step, two molecules of acetyl-coenzyme A are enzymatically combined to form acetoacetyl-CoA. An enzyme known to catalyze this step is, for example, acetyl-CoA thiolase (also known as acetyl-CoA acetyltransferase). Illustrative examples of nucleotide sequences include but are not limited to the following GenBank accession numbers and the organism from which the sequences derived: (NC\_000913 REGION: 2324131..2325315; *Escherichia coli*), (D49362; *Paracoccus denitrificans*), and (L20428; *Saccharomyces cerevisiae*).

In the second step of the MEV pathway, acetoacetyl-CoA is enzymatically condensed with another molecule of acetyl-

CoA to form 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA). An enzyme known to catalyze this step is, for example, HMG-CoA synthase. Illustrative examples of nucleotide sequences include but are not limited to: (NC\_001145 complement 19061..20536; *Saccharomyces cerevisiae*), (X96617; *Saccharomyces cerevisiae*), (X83882; *Arabidopsis thaliana*), (AB037907; *Kitasatospora griseola*), (BT007302; *Homo sapiens*), and (NC\_002758, Locus tag SAV2546, GeneID 1122571; *Staphylococcus aureus*).

10 In the third step, HMG-CoA is enzymatically converted to mevalonate. An enzyme known to catalyze this step is, for example, HMG-CoA reductase. Illustrative examples of nucleotide sequences include but are not limited to: (NM\_206548; *Drosophila melanogaster*), (NC\_002758, Locus tag SAV2545, GeneID 1122570; *Staphylococcus aureus*), (NM\_204485; *Gallus gallus*), (AB015627; *Streptomyces* sp. KO 3988), (AF542543; *Nicotiana attenuata*), (AB037907; *Kitasatospora griseola*), (AX128213, providing the sequence encoding a truncated HMGR; *Saccharomyces cerevisiae*), and (NC\_001145: complement (115734..118898; *Saccharomyces cerevisiae*).

15 In the fourth step, mevalonate is enzymatically phosphorylated to form mevalonate 5-phosphate. An enzyme known to catalyze this step is, for example, mevalonate kinase. Illustrative examples of nucleotide sequences include but are not limited to: (L77688; *Arabidopsis thaliana*), and (X55875; *Saccharomyces cerevisiae*).

20 In the fifth step, a second phosphate group is enzymatically added to mevalonate 5-phosphate to form mevalonate 5-pyrophosphate. An enzyme known to catalyze this step is, for example, phosphomevalonate kinase. Illustrative examples of nucleotide sequences include but are not limited to: (AF429385; *Hevea brasiliensis*), (NM\_006556; *Homo sapiens*), and (NC\_001145. complement 712315..713670; *Saccharomyces cerevisiae*).

25 In the sixth step, mevalonate 5-pyrophosphate is enzymatically converted into IPP. An enzyme known to catalyze this step is, for example, mevalonate pyrophosphate decarboxylase. Illustrative examples of nucleotide sequences include but are not limited to: (X97557; *Saccharomyces cerevisiae*), (AF290095; *Enterococcus faecium*), and (U49260; *Homo sapiens*). DXP Pathway

30 A schematic representation of the DXP pathway is described in FIG. 4. In general, the DXP pathway comprises seven steps. In the first step, pyruvate is condensed with D-glyceraldehyde 3-phosphate to make 1-deoxy-D-xylulose-5-phosphate. An enzyme known to catalyze this step is, for example, 1-deoxy-D-xylulose-5-phosphate synthase. Illustrative examples of nucleotide sequences include but are not limited to: (AF035440; *Escherichia coli*), (NC\_002947, locus tag PP0527; *Pseudomonas putida* KT2440), (CP000026, locus tag SPA2301; *Salmonella enterica Paratyphi*, see ATCC 9150), (NC\_007493, locus tag RSP\_0254; *Rhodobacter sphaeroides* 2.4.1), (NC\_005296, locus tag RPA0952; *Rhodopseudomonas palustris* CGA009), (NC\_004556, locus tag PD1293; *Xylella fastidiosa* Temecula), and (NC\_003076, locus tag AT5G11380; *Arabidopsis thaliana*).

35 In the second step, 1-deoxy-D-xylulose-5-phosphate is converted to 2C-methyl-D-erythritol-4-phosphate. An enzyme known to catalyze this step is, for example, 1-deoxy-D-xylulose-5-phosphate reductoisomerase. Illustrative examples of nucleotide sequences include but are not limited to: (AB013300; *Escherichia coli*), (AF148852; *Arabidopsis thaliana*), (NC\_002947, locus tag PP1597; *Pseudomonas putida* KT2440), (AL939124, locus tag SCO5694; *Strepto-*

13

*myces coelicolor* A3(2)), (NC\_007493, locus tag RSP\_2709; *Rhodobacter sphaeroides* 2.4.1), and (NC\_007492, locus tag Pfl\_1107; *Pseudomonas fluorescens* PfO-1).

In the third step, 2C-methyl-D-erythritol-4-phosphate is converted to 4-diphosphocytidyl-2C-methyl-D-erythritol. An enzyme known to catalyze this step is, for example, 4-diphosphocytidyl-2C-methyl-D-erythritol synthase. Illustrative examples of nucleotide sequences include but are not limited to: (AF230736; *Escherichia coli*), (NC\_007493, locus\_tag RSP\_2835; *Rhodobacter sphaeroides* 2.4.1), (NC\_003071, locus\_tag AT2G02500; *Arabidopsis thaliana*), and (NC\_002947, locus\_tag PP1614; *Pseudomonas putida* KT2440).

In the fourth step, 4-diphosphocytidyl-2C-methyl-D-erythritol is converted to 4-diphosphocytidyl-2C-methyl-D-erythritol-2-phosphate. An enzyme known to catalyze this step is, for example, 4-diphosphocytidyl-2C-methyl-D-erythritol kinase. Illustrative examples of nucleotide sequences include but are not limited to: (AF216300; *Escherichia coli*) and (NC\_007493, locus\_tag RSP\_1779; *Rhodobacter sphaeroides* 2.4.1).

In the fifth step, 4-diphosphocytidyl-2C-methyl-D-erythritol-2-phosphate is converted to 2C-methyl-D-erythritol 2,4-cyclodiphosphate. An enzyme known to catalyze this step is, for example, 2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase. Illustrative examples of nucleotide sequences include but are not limited to: (AF230738; *Escherichia coli*), (NC\_007493, locus\_tag RSP\_6071; *Rhodobacter sphaeroides* 2.4.1), and (NC\_002947, locus\_tag PP1618; *Pseudomonas putida* KT2440).

In the sixth step, 2C-methyl-D-erythritol 2,4-cyclodiphosphate is converted to 1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate. An enzyme known to catalyze this step is, for example, 1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate synthase. Illustrative examples of nucleotide sequences include but are not limited to: (AY033515; *Escherichia coli*), (NC\_002947, locus\_tag PP0853; *Pseudomonas putida* KT2440), and (NC\_007493, locus\_tag RSP\_2982; *Rhodobacter sphaeroides* 2.4.1).

In the seventh step, 1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate is converted into either IPP or its isomer, DMAPP. An enzyme known to catalyze this step is, for example, isopentyl/dimethylallyl diphosphate synthase. Illustrative examples of nucleotide sequences include but are not limited to: (AY062212; *Escherichia coli*) and (NC\_002947, locus\_tag PP0606; *Pseudomonas putida* KT2440).

In some embodiments, “cross talk” (or interference) between the host cell’s own metabolic processes and those processes involved with the production of IPP as provided by the present invention are minimized or eliminated entirely. For example, cross talk is minimized or eliminated entirely when the host microorganism relies exclusively on the DXP pathway for synthesizing IPP, and a MEV pathway is introduced to provide additional IPP. Such host organisms would not be equipped to alter the expression of the MEV pathway enzymes or process the intermediates associated with the MEV pathway. Organisms that rely exclusively or predominantly on the DXP pathway include, for example, *Escherichia coli*.

In some embodiments, the host cell produces IPP via the MEV pathway, either exclusively or in combination with the DXP pathway. In other embodiments, a host’s DXP pathway is functionally disabled so that the host cell produces IPP exclusively through a heterologously introduced MEV pathway. The DXP pathway can be functionally disabled by disabling gene expression or inactivating the function of one or more of the naturally occurring DXP pathway enzymes.

14

In other embodiments, the host cell produces IPP via the DXP pathway, either exclusively or in combination with the MEV pathway. In other embodiments, a host’s MEV pathway is functionally disabled so that the host cell produces IPP exclusively through a heterologously introduced DXP pathway. The MEV pathway can be functionally disabled by disabling gene expression or inactivating the function of one or more of the naturally occurring MEV pathway enzymes.

Methods for genetically modifying host organisms and their cultivation have been previously described. Illustrative examples include U.S. Pat. Nos. 6,689,593; 7,172,886; 7,183,089; U.S. Patent Publication Nos. US 2008/0171378; US 2008/0274523; and US 2009/0203102 and PCT Publication Nos. WO 2007/139924; WO 2009/076676; WO 2010/003007; and WO 2009/132220, which are all incorporated herein by reference in their entirety. Additional methods for modifying host organisms to make isoprene are also provided in the Examples below.

#### Purification and Recovery of Microbiially Derived Gaseous Isoprene

The present invention provides methods for handling a gaseous isoprene composition produced from the microbial host cells. When the resulting gaseous isoprene compositions or the above-described gaseous isoprene compositions are treated with the following methods, then the results are the liquid isoprene compositions described above.

In one aspect, a system for purifying isoprene without extractive distillation is provided. Extractive distillation is defined as distillation in the presence of a solvent that forms no azeotrope with other components in the mixture and is used to separate mixtures that cannot be separated by simple distillation because the volatility of at least two of the components in the mixture is nearly the same, causing them to evaporate at nearly the same temperature at a similar rate. Generally miscible, high boiling, and relatively non-volatile, the extraction distillation solvent interacts differently with the components in the mixture enabling the mixture to be separated by normal distillation. Extractive distillation is almost always used in purifying petroleum-derived isoprene. Because extractive distillation requires special equipment and is inherently energy intensive, it is substantial part of the costs associated with making isoprene. In many embodiments of the present invention, the resulting isoprene compositions do not include trace amounts of an extraction distillation solvent because extractive distillation solvents are not used. Illustrative examples of such solvents include but are not limited to acetonitrile and dimethylformamide.

FIG. 1 is a schematic representation of an exemplary separation system. Host cells are cultivated in a bioreactor and the isoprene produced by the cells vaporizes and forms a gaseous isoprene composition (1). Optionally, the gaseous isoprene composition may pass through a drying process to remove some of the water vapor (not shown). The gaseous isoprene composition (1) is directed to a first chiller 102 which cools the gaseous isoprene composition to a temperature between about 10° C. and about -15° C. The cooled gaseous isoprene composition (7) then passes through drum 104, where the water vapor in the gaseous isoprene composition condenses into a liquid and discharged from the process (5). The exiting gaseous isoprene composition (8) may pass through a drying process to remove any remaining water (not shown). The gaseous isoprene composition (8) is directed to a second chiller 106 further cooling the composition to a temperature below -35° C. The resulting liquid isoprene composition (9) flows to drum 108. Optionally, the bottom stream (10) from drum 108 may then be passed to nitrogen stripper 110 while the top stream (11) is recycled back to the first chiller 102 to

## 15

assist in (or to serve as the refrigerant for) chilling the incoming gaseous isoprene (1) and then exiting a by-product stream (6). Substantially pure nitrogen (2) is introduced into nitrogen stripper 110 whereby the liquid isoprene composition (3) is recovered and the by-product nitrogen gas (4) can be discharged or recovered in a subsequent recovery step (not shown).

In another embodiment, the system comprises:

- a. a bioreactor capable of culturing a plurality of host cells, preferably the bioreactor has a capacity of greater than 100 liters;
- b. a first chiller and flash drum operably connected to the overhead stream of the bioreactor, the first chiller preferably capable of operating in a temperature range of between 10° C. and -15° C.;
- c. a second chiller and flash drum operably connected to the overhead stream exiting from the first chiller and flash drum, the second chiller preferably capable of operating in a temperature below -35° C., for example between about -65° C. and about -85° C.;
- d. optionally, the exiting overhead stream from the second chiller and flash drum may be operably connected to the inlet of the refrigerant or cooling stream of the first chiller; and
- e. optionally, the condensed stream exiting from the second chiller and flash drum may be operably connected to a nitrogen stripper.

In another aspect, method for recovering isoprene using such a system is provided. The method comprises:

- a. obtaining a first gaseous composition comprising isoprene and water wherein the gaseous composition comprises 1 part per million or less of C<sub>2</sub>-C<sub>5</sub> alkynes;
- b. flowing the first gaseous composition through a first chiller wherein the first chiller has a temperature of between about 10° C. and about -15° C. thereby resulting in a second gaseous composition and wherein the second gaseous composition comprises less water than the first gaseous composition, for example the second gas composition comprises about 3% by weight or less of water;
- c. flowing the second gaseous composition through a second chiller wherein the second chiller has a temperature below -35° C.; and
- d. collecting the resulting liquid isoprene composition.

In other embodiments, the method for recovering isoprene comprises reducing the water content present in the first gaseous composition by flowing the first gaseous composition through a first chiller wherein the first chiller has a temperature of between about 10° C. and about -15° C. thereby resulting in a second gaseous composition and wherein the second gaseous composition comprises less water than the first gaseous composition, for example the second gas composition comprises less than about 3% by weight of water. In other embodiments, the second gas composition comprises less than about 2% by weight of water, less than 1%, 0.5%, 0.1%, and 0.05% by weight. In other embodiments, the second gas composition comprises less than about 500 ppm by weight of water, less than 250 ppm, 100 ppm and 50 ppm by weight.

In another aspect, method for recovering isoprene using such a system is provided. The method comprises:

- a. obtaining a first gaseous composition comprising isoprene and water wherein the gaseous composition comprises 1 part per million or less of C<sub>2</sub>-C<sub>5</sub> alkynes;
- b. flowing the first gaseous composition through a first chiller wherein the first chiller has a temperature of between about 10° C. and about -15° C. thereby result-

## 16

ing in a second gaseous composition and wherein the second gaseous composition comprises less water than the first gaseous composition, for example the second gas composition comprises about 3% by weight or less of water;

- c. flowing the second gaseous composition through a second chiller wherein the second chiller has a temperature below -35° C. thereby resulting in a liquid isoprene composition, wherein the liquid isoprene composition comprises less of one or more of the following components than the second gaseous composition, comprising:
  - i. water in amount less than about 1% by weight or less;
  - ii. carbon dioxide in amount less than 1% by weight;
  - iii. nitrogen in amount less than 1% by weight; and
- d. collecting the resulting liquid isoprene composition.

In some embodiments, the method further comprises nitrogen stripping the liquid isoprene composition. This nitrogen stripping may be accomplished by any suitable method including passing a substantially pure nitrogen stream through the liquid isoprene enriched composition. This nitrogen stream serves to further remove dissolved gases (such, as for example oxygen, carbon dioxide, nitrogen, and argon) and/or remaining water. In other embodiments, nitrogen stripping of the liquid isoprene composition may comprise removing: dissolved oxygen to levels less than about 1%, 0.5%, 0.1%, and 0.05% by weight; dissolved carbon dioxide to levels less than about 1%, 0.5%, 0.1%, and 0.05% by weight; dissolved nitrogen to levels less than about 1%, 0.5%, 0.1%, and 0.05% by weight; dissolved argon to levels less than about 1%, 0.5%, 0.1%, and 0.05% by weight; and any remaining water to levels less than about 1%, 0.5%, 0.1%, 0.05% by weight to levels lower than the level of detection.

In some embodiments, the method further comprises extracting hydrocarbon impurities at some point in the purifying isoprene process. This hydrocarbon extraction may be accomplished by passing a portion or all of the first gaseous composition, second gaseous composition and/or the liquid isoprene composition over or through a modified zeolite membrane and/or molecular sieve. The zeolite membrane and/or molecular sieves may be modified to selectively adsorb either isoprene and not the other hydrocarbons present in the treated composition or vice versa (adsorb other hydrocarbons in the treated composition and not isoprene). In some embodiments, the zeolites and/or molecular sieves are modified by carbonization to provide the selected adsorbability. In some embodiments, the zeolites may be L-type, Y-type, ZSM-5, and/or beta-type. In some embodiments a method for enhancing the selectivity of a zeolite by controlled carbonization as detailed in U.S. Pat. No. 7,041,616, which is hereby incorporated in its entirety by reference, may be used.

In other embodiments, the first gaseous composition further comprises carbon dioxide. In still other embodiments, the first gaseous composition further comprises oxygen. In still other embodiments, the first gaseous composition further comprises nitrogen.

In other embodiments, the first gaseous composition comprises 1 part per million or less of C<sub>2</sub>-C<sub>5</sub> alkynes and cyclopentadiene. In further embodiments, the first gaseous composition further comprises 1 part per million or less of C<sub>2</sub>-C<sub>5</sub> alkynes and piperylene. In still other embodiments, the first gaseous composition comprises 1 part per million or less of C<sub>2</sub>-C<sub>5</sub> alkynes and 1,4-pentadiene. In still other embodiments, the first gaseous composition comprises 1 part per million or less of C<sub>2</sub>-C<sub>5</sub> alkynes, cyclopentadiene, piperylene, and 1,4-pentadiene.

In other embodiments, the first gaseous composition is flowed through a drier prior to flowing through the first

chiller. In still other embodiments, the first gaseous composition is flowed through a drier after flowing through the first chiller but prior to flowing through the second chiller.

In other embodiments, the first chiller has a temperature of about -5° C. In other embodiments, the first chiller is cooled using a propylene refrigeration system. In still other embodiments, the first chiller is cooled using an ammonium refrigeration system.

In other embodiments, the second chiller has a temperature less than about -50° C. In other embodiments, the second chiller has a temperature of about -60° C. to about -85° C. In still other embodiments, the second chiller has a temperature of less than about -65° C. In still other embodiments, the second chiller has a temperature between about -35° C. and about -85° C.

In other embodiments, the liquid isoprene composition comprises at least about 70%, 75%, 80%, 85%, and 90% isoprene by weight. In other embodiments, the liquid isoprene comprises at least about 95%, 96%, 97%, 98%, 99% and 99.5% isoprene by weight. In still other embodiments, the liquid isoprene composition comprises isoprene in an amount that is greater than about 99.5% by weight.

In other embodiments, the second chiller is cooled using an ethylene refrigeration system.

In another aspect, another method is provided. The method comprises:

- a. culturing a plurality of host cells capable of making isoprene;
- b. forming a first gaseous composition comprising isoprene and water wherein the water is present in an amount greater than about 70% of its saturation amount;
- c. subjecting the first gaseous composition to a first cooling step whereby substantially all of the water is removed from the first gaseous composition resulting a second gaseous composition;
- d. subjecting the second gaseous composition to a second cooling step whereby a liquid isoprene composition is collected.

In some embodiments, the method further comprises nitrogen stripping the liquid isoprene composition.

In other embodiments, the first gaseous composition further comprises carbon dioxide. In still other embodiments, the first gaseous composition further comprises oxygen. In still other embodiments, the first gaseous composition further comprises nitrogen.

In other embodiments, the first gaseous composition comprises 1 part per million or less of C<sub>2</sub>-C<sub>5</sub> alkynes and cyclopentadiene. In further embodiments, the first gaseous composition further comprises 1 part per million or less of C<sub>2</sub>-C<sub>5</sub> alkynes and piperylene. In still other embodiments, the first gaseous composition comprises 1 part per million or less of C<sub>2</sub>-C<sub>5</sub> alkynes and 1,4-pentadiene. In still other embodiments, the first gaseous composition comprises 1 part per million or less of C<sub>2</sub>-C<sub>5</sub> alkynes, cyclopentadiene, piperylene, and 1,4-pentadiene.

In other embodiments, the first gaseous composition is subjected through a drier prior to the first cooling step. In still other embodiments, the first gaseous composition is subjected through a drier after the first cooling step but prior to the second cooling step.

In other embodiments, the first cooling step cools the first gaseous isoprene composition to a temperature between about 10° C. and about -15° C., between about 10° C. and about -10° C., about 5° C. and about -5° C., and about 5° C. and about -10° C.

In other embodiments, the first cooling step uses a propylene refrigeration system. In still other embodiments, the first cooling step uses chiller an ammonium refrigeration system.

In other embodiments, the second cooling step cools the second gaseous isoprene composition to a temperature less than -35° C. In other embodiments, the second cooling step cools the second gaseous isoprene composition to a temperature less than about -50° C. In still other embodiments, the second cooling step cools the second gaseous isoprene composition to a temperature about -60° C. to about -85° C. In still further embodiments, the second cooling step cools the second gaseous isoprene composition to a temperature of less than about -65° C.

In other embodiments, the second cooling step uses an ethylene refrigeration system.

In other embodiments, the liquid isoprene composition comprises at least about 70%, 75%, 80%, 85%, and 90% isoprene by weight. In other embodiments, the liquid isoprene comprises at least about 95%, 96%, 97%, 98%, 99% and 99.5% isoprene by weight. In still other embodiments, the liquid isoprene composition comprises isoprene in an amount that is greater than about 99.5% by weight.

In other embodiments, the host cells are selected from the genus *Bacillus*, *Escherichia* or *Acinetobacter*. In still other embodiments, the host cells are *Escherichia coli*. In further embodiments, the host cells are yeast. In still further embodiments, the host cells are *Saccharomyces cerevisiae*.

In another aspect, another method is provided. The method comprises:

- a. contacting a plurality of host cells capable of making isoprene in an aqueous medium wherein the aqueous medium is in contact with an immiscible organic liquid and the aqueous medium, the host cells, and the immiscible organic liquid is in a closed vessel; and
- b. culturing the host cells in the aqueous medium whereby the host cells make isoprene and the isoprene is captured in the immiscible organic liquid.

In some embodiments, the method further comprises separating the immiscible organic liquid from the aqueous medium and separating the isoprene from the immiscible organic liquid.

In other embodiments, the immiscible organic liquid is selected from butyl acetate, ethyl acetate, isopropyl myristate, methyl isobutyl ketone, methyl oleate, and toluene. In certain embodiments, the solvent is butyl acetate. In other embodiments, the immiscible organic liquid is isopropyl myristate.

In other embodiments, the isoprene is separated from the immiscible organic liquid by heating the immiscible organic liquid to a temperature above 34° C.

The resulting gaseous isoprene composition can then be further purified using the methods described above.

In another aspect, a system for making microbial isoprene is provided. An illustrative example of such a system is shown in FIG. 2. Bioreactor 202 is a closed system where host cells capable of making isoprene are cultivated in an aqueous medium and with an immiscible organic liquid on top of the aqueous medium. Because bioreactor 202 is a closed system, the isoprene produced by the host cells is captured by the immiscible organic liquid. The isoprene-enriched immiscible organic liquid (A) may then directed be directed to any suitable gas-liquid separation method. In this example, the isoprene-enriched immiscible organic liquid is directed to a heater 204 which volatizes the isoprene into a gaseous isoprene composition (1). This gaseous isoprene composition (1) can then be further purified using the systems and methods

19

described above. The immiscible organic liquid (B) can optionally be recycled and used in subsequent bioreactions to make isoprene.

In another embodiment, the system comprises:

- a. a closed vessel;
- b. an aqueous medium, within the vessel, forming a first phase;
- c. a plurality of host cells, within the aqueous medium, capable of making isoprene; and,
- d. a liquid organic second phase, capable of capturing the isoprene made by the host cells, in contact with the first phase.

In some embodiments, the immiscible organic liquid is selected from butyl acetate, ethyl acetate, isopropyl myristate, methyl isobutyl ketone, methyl oleate, and toluene. In certain embodiments, the solvent is butyl acetate. In other embodiments, the immiscible organic liquid is isopropyl myristate.

In other embodiments, the host cells are selected from the genus *Bacillus*, *Escherichia* or *Acinetobacter*. In still other embodiments, the host cells are *Escherichia coli*. In further embodiments, the host cells are yeast. In still further embodiments, the host cells are *Saccharomyces cerevisiae*.

#### EXAMPLES

The practice of the present invention can employ, unless otherwise indicated, conventional techniques of the biosynthetic industry and the like, which are within the skill of the art. To the extent such techniques are not described fully herein, one can find ample reference to them in the scientific literature.

In the following examples, efforts have been made to ensure accuracy with respect to numbers used (for example, amounts, temperature, and so on), but variation and deviation can be accommodated, and in the event a clerical error in the numbers reported herein exists, one of ordinary skill in the arts to which this invention pertains can deduce the correct amount in view of the remaining disclosure herein. Unless indicated otherwise, temperature is reported in degrees Celsius, and pressure is at or near atmospheric pressure at sea level. All reagents, unless otherwise indicated, were obtained commercially. The following examples are intended for illustrative purposes only and do not limit in any way the scope of the present invention.

#### Example 1

This example describes methods for detecting isoprene trapped in an immiscible organic liquid in a closed fermentation system.

In this example, isopropyl myristate (IPM) and solutions of IPM spiked with 103 mg/L, 10.3 mg/L and 0 mg/L isoprene were prepared and stored in 100 ml capped media bottles.

125 ml unbaffled flasks with screw caps/septa were set up in triplicate. The flasks contained 40 ml medium (Yeast Nitrogen Base media with 4% galactose, 0.2% glucose, Leu) inoculated with overnight yeast culture (which does not make isoprene) grown to an OD=0.05 with 8 ml of IPM solutions with the various concentrations of isoprene. The sealed flasks were incubated for 72 hours at 30° C. and 200 rpm.

Post-incubation, the IPM overlay was phase-separated by manual transfer to primary GC vials, then transferred to secondary GC vials and run undiluted on GC-FID. In addition to samples from flasks, the original solutions of 103 and 10.3 mg/L isoprene in IPM (NOT shaken for 72 hours at 200 rpm, 30° C.) were also analyzed.

20

As shown by FIG. 5, over 90% of the spiked isoprene was recovered from the IPM layer. Cell growth for cultures with 103, 10.3, and 0 mg/L isoprene in the IPM layer was indistinguishable.

5

#### Example 2

This example describes methods for making nucleic acids for expressing in *Saccharomyces cerevisiae* heterologous isoprene synthases.

Genomic DNA was isolated from *Saccharomyces cerevisiae* strains Y002 (CEN.PK2 background MATA ura3-52 trp1-289 leu2-3,112 his3Δ1 MAL2-8C SUC2) (van Dijken et al. (2000) Enzyme and Microbial Technology 26:706-714), 15 Y007 (S288C background MATA trp1Δ63) (ATCC number 200873), Y051 (S288C background), and EG123 (ATCC number 204278). The strains were grown overnight in liquid medium containing 1% Yeast extract, 2% Bacto-peptone, and 20 2% Dextrose (YPD medium). Cells were isolated from 10 mL liquid cultures by centrifugation at 3,100 rpm, washing of cell pellets in 10 mL ultra-pure water, and re-centrifugation. Genomic DNA was extracted using the Y-DER yeast DNA extraction kit (Pierce Biotechnologies, Rockford, Ill.) as per manufacturer's suggested protocol. Extracted genomic DNA 25 was re-suspended in 100 uL 10 mM Tris-Cl, pH 8.5, and OD<sub>260/280</sub> readings were taken on a ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, Del.) to determine genomic DNA concentration and purity.

DNA amplification by Polymerase Chain Reaction (PCR) 30 was done in an Applied Biosystems 2720 Thermocycler (Applied Biosystems Inc, Foster City, Calif.) using the Phusion High Fidelity DNA Polymerase system (Finnzymes OY, Espoo, Finland) as per manufacturer's suggested protocol. Upon the completion of a PCR amplification of a DNA fragment 35 that was to be inserted into the pCR®4Blunt-TOPO vector (Invitrogen, Carlsbad, Calif.), A nucleotide overhangs were created by adding 1 uL of Qiagen Taq Polymerase (Qiagen, Valencia, Calif.) to the reaction mixture and performing an additional 10 minute, 72° C. PCR extension step, 40 followed by cooling to 4° C. Upon completion of a PCR amplification, 8 uL of a 50% glycerol solution was added to the reaction mix.

Agarose gel electrophoresis was performed using a 1% TBE (0.89 M Tris, 0.89 M Boric acid, 0.02 M EDTA sodium 45 salt) agarose gel containing 0.5 μg/mL ethidium bromide, at 120 V, 400 mA for 30 minutes. DNA bands were visualized using ultraviolet light. DNA bands were excised from the gel with a sterile razor blade, and the excised DNA was gel purified using the Zymoclean Gel DNA Recovery Kit (Zymo Research, Orange, Calif.) according to manufacturer's suggested protocols. The purified DNA was eluted into 10 uL ultra-pure water, and OD<sub>260/280</sub> readings were taken on a ND-1000 spectrophotometer to determine DNA concentration and purity.

Ligations were performed using 100-500 μg of purified 50 PCR product and High Concentration T4 DNA Ligase (New England Biolabs, Ipswich, Mass.) as per manufacturer's suggested protocol. For plasmid propagation, ligated constructs were transformed into *Escherichia coli* DH5α chemically competent cells (Invitrogen, Carlsbad, Calif.) as per manufacturer's suggested protocol. Positive transformants were selected on solid media containing 1.5% Bacto Agar, 1% Tryptone, 0.5% Yeast Extract, 1% NaCl, and an appropriate antibiotic. Isolated transformants were grown for 16 hours in liquid LB medium containing 50 μg/mL carbenicillin or kanamycin antibiotic at 37° C., and plasmid was isolated and purified using a QIAprep Spin Miniprep kit (Qiagen, Valen-

cia, Calif.) as per manufacturer's suggested protocol. Constructs were verified by performing diagnostic restriction enzyme digestions, resolving DNA fragments on an agarose gel, and visualizing the bands using ultraviolet light. Select constructs were also verified by DNA sequencing, which was done by Elim Biopharmaceuticals Inc. (Hayward, Calif.).

Expression plasmid pAM353 was generated by inserting a nucleotide sequence encoding a  $\beta$ -farnesene synthase into the pRS425-Gal1 vector (Mumberg et. al. (1994) *Nucl. Acids. Res.* 22(25): 5767-5768). The nucleotide sequence insert was generated synthetically, using as a template the coding sequence of the  $\beta$ -farnesene synthase gene of *Artemisia annua* (GenBank accession number AY835398) codon-optimized for expression in *Saccharomyces cerevisiae* (SEQ ID NO: 1). The synthetically generated nucleotide sequence was flanked by 5' BamHI and 3' XhoI restriction sites, and could thus be cloned into compatible restriction sites of a cloning vector such as a standard pUC or pACYC origin vector. The synthetically generated nucleotide sequence was isolated by digesting to completion the construct using BamHII and XhoI restriction endonucleases. The reaction mixture was resolved by gel electrophoresis, the approximately 1.7 kb DNA fragment comprising the  $\beta$ -farnesene synthase coding sequence was gel purified, and the isolated DNA fragment was ligated into the BamHI XhoI restriction site of the pRS425-Gal1 vector, yielding expression plasmid pAM353.

Expression plasmid pAM404 was generated by inserting a nucleotide sequence encoding the  $\beta$ -farnesene synthase of *Artemisia annua* (GenBank accession number AY835398) codon-optimized for expression in *Saccharomyces cerevisiae* into vector pAM178 (SEQ ID NO: 2). The nucleotide sequence encoding the  $\beta$ -farnesene synthase was PCR amplified from pAM353 using primers 52-84 pAM326 BamHI (SEQ ID NO: 21) and 52-84 pAM326 NheI (SEQ ID NO: 22). The resulting PCR product was digested to completion using BamHII and NheI restriction endonucleases, the reaction mixture was resolved by gel electrophoresis, the approximately 1.7 kb DNA fragment comprising the  $\beta$ -farnesene synthase coding sequence was gel purified, and the isolated DNA fragment was ligated into the BamHI NheI restriction site of vector pAM178, yielding expression plasmid pAM404.

Plasmid Genetrix2080 was generated by inserting a nucleotide sequence encoding an isoprene synthase into vector pUC19. The insert was generated synthetically as two approximately equal sized DNA fragments, fragment 2080\_1 (SEQ ID NO: 3) and fragment 2080\_2 (SEQ ID NO: 4), using as a template the coding sequence of the isoprene synthase gene of Kudzu codon-optimized for expression in *Saccharomyces cerevisiae*. Each DNA fragment was flanked by Lgul restriction sites, and comprised a 40 base pair overlapping sequence at one end. The synthetically generated DNA fragments were blunt ligated into the SmaI restriction site of the pUC19 cloning vector, from which the two inserts were excised again by digesting to completion 500  $\mu$ g of the construct using Lgul restriction endonuclease (Fermentas, Glen Burnie, Md.). The restriction endonuclease was heat inactivated for 20 minutes at 65° C., and the DNA fragments were stitched together by a first round of PCR amplification (one cycle of denature at 98° C. for 2 minutes; 5 cycles of denature at 98° C. for 30 seconds and anneal/extend at 72° C. for 30 seconds per kilobase PCR product; no primers were used). Samples were placed on ice, 0.5 uM of each terminal primer TRIX\_L\_494 (SEQ ID NO: 79) and TRIX\_L\_495 (SEQ ID NO: 80) were added to the reaction mixture, and a second round of PCR amplification was performed (one cycle of denature at 98° C. for 2 minutes; 35 rounds of denature at 98° C. for 12 seconds and anneal/extend at 72° C. for 20

seconds per kilobase PCR product; one cycle of final extend at 72° C. for 7 minutes; and a final hold at 4° C.). The reaction mixture was resolved by gel electrophoresis, the assembled DNA fragment was gel purified, treated with T4 polynucleotide kinase (PNK) (New England Biolabs, Ipswich, Mass.), and blunt ligated into the SmaI restriction site of vector pUC19, yielding plasmid Genetrix2080.

Plasmid Genetrix2081 was generated by inserting a nucleotide sequence encoding an isoprene synthase into vector pUC19. The insert was generated synthetically as two approximately equal sized DNA fragments, fragment 2081\_1 (SEQ ID NO: 5) and fragment 2081\_2 (SEQ ID NO: 6), using as a template the coding sequence of the isoprene synthase gene of *Populus nigra* codon-optimized for expression in *Saccharomyces cerevisiae*. Each DNA fragment was flanked by Lgul restriction sites, and comprised a 30 base pair overlapping sequence at one end. The synthetically generated DNA fragments were blunt ligated into the SmaI restriction site of the pUC19 cloning vector, from which the two inserts were excised again by digesting to completion 500  $\mu$ g of the construct using Lgul restriction endonuclease (Fermentas, Glen Burnie, Md.). The restriction endonuclease was heat inactivated for 20 minutes at 65° C., and the DNA fragments were stitched together by a first round of PCR amplification (one cycle of denature at 98° C. for 2 minutes; 5 cycles of denature at 98° C. for 30 seconds and anneal/extend at 72° C. for 30 seconds per kilobase PCR product; no primers were used). Samples were placed on ice, 0.5 uM of each terminal primer TRIX\_L\_497 (SEQ ID NO: 81) and TRIX\_L\_498 (SEQ ID NO: 82) were added to the reaction mixture, and a second round of PCR amplification was performed (one cycle of denature at 98° C. for 2 minutes; 35 rounds of denature at 98° C. for 12 seconds and anneal/extend at 72° C. for 20 seconds per kilobase PCR product; one cycle of final extend at 72° C. for 7 minutes; and a final hold at 4° C.). The reaction mixture was resolved by gel electrophoresis, the assembled DNA fragment was gel purified, treated with T4 polynucleotide kinase (PNK) (New England Biolabs, Ipswich, Mass.), and blunt ligated into the SmaI restriction site of vector pUC19, yielding plasmid Genetrix2081.

Plasmid Genetrix2082 was generated by inserting a nucleotide sequence encoding an isoprene synthase into vector pUC19. The insert was generated synthetically as two approximately equal sized DNA fragments, fragment 2082\_1 (SEQ ID NO: 7) and fragment 2082\_2 (SEQ ID NO: 8), using as a template the coding sequence of the isoprene synthase gene of *Populus alba*/*Populus tremula* codon-optimized for expression in *Saccharomyces cerevisiae*. Each DNA fragment was flanked by Lgul restriction sites, and comprised a 40 base pair overlapping sequence at one end. The synthetically generated DNA fragments were blunt ligated into the SmaI restriction site of the pUC19 cloning vector, from which the two inserts were excised again by digesting to completion 500  $\mu$ g of the construct using Lgul restriction endonuclease (Fermentas, Glen Burnie, Md.). The restriction endonuclease was heat inactivated for 20 minutes at 65° C., and the DNA fragments were stitched together by a first round of PCR amplification (one cycle of denature at 98° C. for 2 minutes; 5 cycles of denature at 98° C. for 30 seconds and anneal/extend at 72° C. for 30 seconds per kilobase PCR product; no primers were used). Samples were placed on ice, 0.5 uM of each terminal primer TRIX\_L\_500 (SEQ ID NO: 83) and TRIX\_L\_501 (SEQ ID NO: 84) were added to the reaction mixture, and a second round of PCR amplification was performed (one cycle of denature at 98° C. for 2 minutes; 35 rounds of denature at 98° C. for 12 seconds and anneal/extend at 72° C. for 20 seconds per kilobase PCR product; one

23

cycle of final extend at 72° C. for 7 minutes; and a final hold at 4° C.). The reaction mixture was resolved by gel electrophoresis, the assembled DNA fragment was gel purified, treated with T4 polynucleotide kinase (PNK) (New England Biolabs, Ipswich, Mass.), and blunt ligated into the SmaI restriction site of vector pUC19, yielding plasmid Genetrix2082.

Expression plasmid pAM1547 was generated by replacing the β-farnesene synthase coding sequence of expression plasmid pAM404 with the isoprene synthase coding sequence of plasmid Genetrix2080. DNA fragment IS2080 was generated by PCR amplifying plasmid Genetrix2080 using primers YD-198-70A (SEQ ID NO: 85) and YD-198-70B (SEQ ID NO: 86), digesting the PCR product to completion using restriction endonucleases BamHI and NheI, resolving the reaction mixture by gel electrophoresis, and gel purifying the approximately 1.8 kb DNA fragment comprising the isoprene synthase coding sequence. Expression plasmid pAM404 was digested to completion using restriction endonucleases BamHI and NheI, the reaction mixture was resolved by gel electrophoresis, the approximately 7.3 kb vector backbone (lacking the β-farnesene synthase coding sequence) was gel purified, and the purified vector backbone was ligated with DNA fragment IS2080, yielding pAM1547 (see FIG. 6 for a plasmid map).

Expression plasmid pAM1548 was generated by replacing the β-farnesene synthase coding sequence of expression plasmid pAM404 with a truncated version of the isoprene synthase coding sequence of plasmid Genetrix2080. DNA fragment IS2080T was generated by PCR amplifying plasmid Genetrix2080 using primers YD-198-70B (SEQ ID NO: 86) and YD-198-70G (SEQ ID NO: 91), digesting the PCR product to completion using restriction endonucleases BamHI and NheI, resolving the reaction mixture by gel electrophoresis, and gel purifying the approximately 1.65 kb DNA fragment comprising the truncated isoprene synthase coding sequence. pAM404 was digested to completion using restriction endonucleases BamHI and NheI, the reaction mixture was resolved by gel electrophoresis, the approximately 7.3 kb backbone (lacking the β-farnesene synthase coding sequence) was gel purified, and the purified vector backbone was ligated with the amplified DNA fragment IS2080T, yielding pAM1548.

Expression plasmid pAM1549 was generated by replacing the β-farnesene synthase coding sequence of expression plasmid pAM404 with the isoprene synthase coding sequence of plasmid Genetrix2081. DNA fragment IS2081 was generated by PCR amplifying plasmid Genetrix2081 using primers YD-198-70C (SEQ ID NO: 87) and YD-198-70D (SEQ ID NO: 88), digesting the PCR product to completion using restriction endonucleases BamHI and NheI, resolving the reaction mixture by gel electrophoresis, and gel purifying the approximately 1.8 kb DNA fragment comprising the isoprene synthase coding sequence. Expression plasmid pAM404 was digested to completion using restriction endonucleases BamHI and NheI, the reaction mixture was resolved by gel electrophoresis, the approximately 7.3 kb vector backbone (lacking the β-farnesene synthase coding sequence) was gel purified, and the purified vector backbone was ligated with DNA fragment IS2081, yielding pAM1549.

Expression plasmid pAM1550 was generated by replacing the β-farnesene synthase coding sequence of expression plasmid pAM404 with a truncated version of the isoprene synthase coding sequence of plasmid Genetrix2081. DNA fragment IS2081T was generated by PCR amplifying plasmid Genetrix2081 using primers YD-198-70D (SEQ ID NO: 88) and YD-198-70H (SEQ ID NO: 92), digesting the PCR prod-

24

uct to completion using restriction endonucleases BamHI and NheI, resolving the reaction mixture by gel electrophoresis, and gel purifying the approximately 1.6 kb DNA fragment comprising the truncated isoprene synthase coding sequence.

5 Expression plasmid pAM404 was digested to completion using restriction endonucleases BamHI and NheI, the reaction mixture was resolved by gel electrophoresis, the approximately 7.3 kb vector backbone (lacking the β-farnesene synthase coding sequence) was gel purified, and the purified vector backbone was ligated with DNA fragment IS2081T, yielding pAM1550.

Expression plasmid pAM1551 was generated by replacing the β-farnesene synthase coding sequence of expression plasmid pAM404 with the isoprene synthase coding sequence of plasmid Genetrix2082. DNA fragment IS2082 was generated by PCR amplifying plasmid Genetrix2082 using primers YD-198-70E (SEQ ID NO: 89) and YD-198-70F (SEQ ID NO: 90), digesting the PCR product to completion using restriction endonucleases BamHI and NheI, resolving the reaction mixture by gel electrophoresis, and gel purifying the approximately 1.8 kb DNA fragment comprising the isoprene synthase coding sequence. Expression plasmid pAM404 was digested to completion using restriction endonucleases 10 BamHI and NheI, the reaction mixture was resolved by gel electrophoresis, the approximately 7.3 kb vector backbone (lacking the β-farnesene synthase coding sequence) was gel purified, and the purified vector backbone was ligated with DNA fragment IS2082, yielding pAM1551.

Expression plasmid pAM1552 was generated by replacing the β-farnesene synthase coding sequence of expression plasmid pAM404 with a truncated version of the isoprene synthase coding sequence of plasmid Genetrix2082. DNA fragment IS2082T was generated by PCR amplifying plasmid 15 Genetrix2082 using primers YD-198-70F (SEQ ID NO: 90) and YD-198-70I (SEQ ID NO: 93), digesting the PCR product to completion using restriction endonucleases BamHI and NheI, resolving the reaction mixture by gel electrophoresis, and gel purifying the approximately 1.64 kb DNA fragment comprising the truncated isoprene synthase coding sequence. Expression plasmid pAM404 was digested to completion using restriction endonucleases BamHI and NheI, the reaction mixture was resolved by gel electrophoresis, the approximately 7.3 kb vector backbone (lacking the β-farnesene synthase coding sequence) was gel purified, and the purified vector backbone was ligated with DNA fragment IS2082T, yielding pAM1552.

Plasmid pAM840 was generated by inserting the coding sequence of the hisG gene into the pCR®2.1-TOPO vector 20 (Invitrogen, Carlsbad, Calif.). The coding sequence of the hisG gene was PCR amplified using primers KB34 (SEQ ID NO: 75) and KB39 (SEQ ID NO: 76) and plasmid pNKY51 (Alani et al. (1987) Genetics 116(4):541-555) as template. The amplified DNA fragment was ligated with the Topo vector 25 as per manufacturer's suggested protocol, yielding pAM840.

Plasmid pAM728 was generated by introducing the coding sequence of the farnesene synthase gene of *Artemisia annua* (GenBank accession number AY835398) codon-optimized 30 for expression in *Saccharomyces cerevisiae* and under control of the promoter of the GAL7 gene of *Saccharomyces cerevisiae* (P<sub>GAL7</sub>) into plasmid pRS425 (Christianson et al. (1992) Gene 110(1):119-122). An approximately 0.5 kb DNA fragment comprising P<sub>GAL7</sub> was PCR amplified from Y002 35 genomic DNA using primers GW-110-26-pGAL7-PstI F (SEQ ID NO: 119) and GW-110-26-pGAL7 R (SEQ ID NO: 120) and was gel purified. An approximately 2 kb DNA

fragment comprising the coding sequence of the farnesene synthase gene was PCR amplified using primers GW-110-26-pGAL7-FS F (SEQ ID NO: 121) and GW-110-26-FS-BamHI R (SEQ ID NO: 122). The two DNA fragments were stitched together using PCR primers GW-110-26-pGAL7-PstI F (SEQ ID NO: 119) and GW-110-26-FS-BamHI R (SEQ ID NO: 122) to create a P<sub>GAL7</sub>-FS-tCYC1 insert. The P<sub>GAL7</sub>-FS-tCYC1 insert and plasmid pRS425 were digested to completion using PstI and BamHI restriction endonucleases, and the two DNA fragments were ligated, yielding pAM728.

were generated by PCR amplification as outlined in Table 1. For the construction of pAM489, 400 ng of pAM471 and 100 ng of pAM466 were digested to completion using XmaI restriction enzyme (New England Biolabs, Ipswich, Mass.), DNA fragments corresponding to the ERG20-P<sub>GAL</sub>-tHMGR insert and the linearized pAM466 vector were gel purified, and 4 molar equivalents of the purified insert was ligated with 1 molar equivalent of the purified linearized vector, yielding pAM489. FIG. 7A shows a map of the ERG20-P<sub>GAL</sub>-tHMGR insert, and SEQ ID NO: 9 shows the nucleotide sequence of the DNA fragment and the flanking TRP1 sequences.

TABLE 1

| PCR amplifications performed to generate pAM489 |  |                                   |                                   |                               |
|---|--|-----------------------------------|-----------------------------------|-------------------------------|
| PCR Round                                       | Template   | Primer 1                          | Primer 2                          | PCR Product                   |
| 1   | 100 ng of Y051 genomic DNA   | 61-67-CPK001-G<br>(SEQ ID NO: 23) | 61-67-CPK002-G<br>(SEQ ID NO: 24) | TRP1 <sup>-856 to -226</sup>  |
|   | 100 ng of EG123 genomic DNA  | 61-67-CPK003-G<br>(SEQ ID NO: 25) | 61-67-CPK004-G<br>(SEQ ID NO: 26) | TRP1 <sup>-225 to +548</sup>  |
|   | 100 ng of Y002 genomic DNA   | 61-67-CPK025-G<br>(SEQ ID NO: 47) | 61-67-CPK050-G<br>(SEQ ID NO: 55) | ERG20                         |
|   | 100 ng each of TRP1 <sup>-856 to -226</sup> and TRP1 <sup>-225 to +548</sup> purified PCR products | 61-67-CPK051-G<br>(SEQ ID NO: 56) | 61-67-CPK052-G<br>(SEQ ID NO: 57) | P <sub>GAL</sub>              |
|   | 100 ng each of ERG20 and P <sub>GAL</sub> purified PCR products                                    | 61-67-CPK053-G<br>(SEQ ID NO: 58) | 61-67-CPK031-G<br>(SEQ ID NO: 48) | tHMGR                         |
|   | 100 ng each of ERG20-P <sub>GAL</sub> and tHMGR purified PCR products                              | 61-67-CPK001-G<br>(SEQ ID NO: 23) | 61-67-CPK004-G<br>(SEQ ID NO: 26) | TRP1 <sup>-856 to +548</sup>  |
|   |  | 61-67-CPK025-G<br>(SEQ ID NO: 47) | 61-67-CPK052-G<br>(SEQ ID NO: 57) | ERG20-P <sub>GAL</sub>        |
|   |  | 61-67-CPK025-G<br>(SEQ ID NO: 47) | 61-67-CPK031-G<br>(SEQ ID NO: 48) | ERG20-P <sub>GAL</sub> -tHMGR |

Plasmid pAM940 was generated by introducing the farnesene synthase sequence of plasmid pAM728 into plasmid pRS426 (Christianson et al. (1992) Gene 110(1):119-122). Plasmids pAM728 and pRS426 were digested to completion using XbaI and BamHI restriction endonucleases, the reaction mixtures were resolved by gel electrophoresis, the approximately 5.7 kb pRS426 vector backbone and the approximately 2.5 kb P<sub>GAL7</sub>-FS-tCYC1 insert of pAM728 were gel purified, and the two DNA fragments were ligated, yielding plasmid pAM940.

Plasmid pAM489 was generated by inserting the ERG20-P<sub>GAL</sub>-tHMGR insert of vector pAM471 into vector pAM466. Vector pAM471 was generated by inserting DNA fragment ERG20-P<sub>GAL</sub>-tHMGR, which comprises the open reading frame (ORF) of the ERG20 gene of *Saccharomyces cerevisiae* (ERG20 nucleotide positions 1 to 1208; A of ATG start codon is nucleotide 1) (ERG20), the genomic locus containing the divergent GAL1 and GAL10 promoter of *Saccharomyces cerevisiae* (GAL1 nucleotide position -1 to -668) (P<sub>GAL</sub>), and a truncated ORF of the HMG1 gene of *Saccharomyces cerevisiae* (HMG1 nucleotide positions 1586 to 3323) (tHMGR), into the TOPO Zero Blunt II cloning vector (Invitrogen, Carlsbad, Calif.). Vector pAM466 was generated by inserting DNA fragment TRP1<sup>-856 to +548</sup>, which comprises a segment of the wild-type TRP1 locus of *Saccharomyces cerevisiae* that extends from nucleotide position -856 to position 548 and harbors a non-native internal XmaI restriction site between bases -226 and -225, into the TOPO TA pCR2.1 cloning vector (Invitrogen, Carlsbad, Calif.). DNA fragments ERG20-P<sub>GAL</sub>-tHMGR and TRP1<sup>-856 to +548</sup>

Plasmid pAM491 was generated by inserting the ERG13-P<sub>GAL</sub>-tHMGR insert of vector pAM472 into vector pAM467. Vector pAM472 was generated by inserting DNA fragment ERG13-P<sub>GAL</sub>-tHMGR, which comprises the ORF of the ERG13 gene of *Saccharomyces cerevisiae* (ERG13 nucleotide positions 1 to 1626) (ERG13), the genomic locus containing the divergent GAL1 and GAL10 promoter of *Saccharomyces cerevisiae* (GAL1 nucleotide position -1 to -668) (P<sub>GAL</sub>), and a truncated ORF of the HMG1 gene of *Saccharomyces cerevisiae* (HMG1 nucleotide position 1586 to 3323) (tHMGR), into the TOPO Zero Blunt II cloning vector. Vector pAM467 was generated by inserting DNA fragment URA3<sup>-723 to 701</sup>, which comprises a segment of the wild-type URA3 locus of *Saccharomyces cerevisiae* that extends from nucleotide position -723 to position -224 and harbors a non-native internal XmaI restriction site between bases -224 and -223, into the TOPO TA pCR2.1 cloning vector. DNA fragments ERG13-P<sub>GAL</sub>-tHMGR and URA3<sup>-723 to 701</sup> were generated by PCR amplification as outlined in Table 2. For the construction of pAM491, 400 ng of pAM472 and 100 ng of pAM467 were digested to completion using XmaI restriction enzyme, DNA fragments corresponding to the ERG13-P<sub>GAL</sub>-tHMGR insert and the linearized pAM467 vector were gel purified, and 4 molar equivalents of the purified insert was ligated with 1 molar equivalent of the purified linearized vector, yielding pAM491. FIG. 7B shows a map of the ERG13-P<sub>GAL</sub>-tHMGR insert, and SEQ ID NO: 10 shows the nucleotide sequence of the DNA fragment and the flanking URA3 sequences.

TABLE 2

| PCR amplifications performed to generate pAM491 |   |                                   |                                   |                               |
|---|---|-----------------------------------|-----------------------------------|-------------------------------|
| PCR Round                                       | Template  | Primer 1                          | Primer 2                          | PCR Product                   |
| 1   | 100 ng of Y007 genomic DNA  | 61-67-CPK005-G<br>(SEQ ID NO: 47) | 61-67-CPK006-G<br>(SEQ ID NO: 28) | URA3 <sup>-723 to -224</sup>  |
|   |   | 61-67-CPK007-G<br>(SEQ ID NO: 29) | 61-67-CPK008-G<br>(SEQ ID NO: 30) | URA3 <sup>-223 to 701</sup>   |
|   | 100 ng of Y002 genomic DNA  | 61-67-CPK032-G<br>(SEQ ID NO: 49) | 61-67-CPK054-G<br>(SEQ ID NO: 59) | ERG13                         |
|   |   | 61-67-CPK052-G<br>(SEQ ID NO: 57) | 61-67-CPK055-G<br>(SEQ ID NO: 60) | P <sub>GAL</sub>              |
|   |   | 61-67-CPK031-G<br>(SEQ ID NO: 48) | 61-67-CPK053-G<br>(SEQ ID NO: 58) | tHMGR                         |
| 2   | 100 ng each of URA3 <sup>-723 to -224</sup> and URA3 <sup>-223 to 701</sup> purified PCR products | 61-67-CPK005-G<br>(SEQ ID NO: 27) | 61-67-CPK008-G<br>(SEQ ID NO: 30) | URA3 <sup>-723 to 701</sup>   |
|   | 100 ng each of ERG13 and P <sub>GAL</sub> purified PCR products                                   | 61-67-CPK032-G<br>(SEQ ID NO: 49) | 61-67-CPK052-G<br>(SEQ ID NO: 57) | ERG13-P <sub>GAL</sub>        |
| 3   | 100 ng each of ERG13-P <sub>GAL</sub> and tHMGR purified PCR products                             | 61-67-CPK031-G<br>(SEQ ID NO: 48) | 61-67-CPK032-G<br>(SEQ ID NO: 49) | ERG13-P <sub>GAL</sub> -tHMGR |

Plasmid pAM493 was generated by inserting the IDI1-P<sub>GAL</sub>-tHMGR insert of vector pAM473 into vector pAM468. Vector pAM473 was generated by inserting DNA fragment IDI1-P<sub>GAL</sub>-tHMGR, which comprises the ORF of the IDI1 gene of *Saccharomyces cerevisiae* (IDI1 nucleotide position 1 to 1017) (IDI1), the genomic locus containing the divergent GAL1 and GAL10 promoter of *Saccharomyces cerevisiae*

and the linearized pAM468 vector were gel purified, and 4 molar equivalents of the purified insert was ligated with 1 molar equivalent of the purified linearized vector, yielding vector pAM493. FIG. 7C shows a map of the IDI1-P<sub>GAL</sub>-tHMGR insert, and SEQ ID NO: 11 shows the nucleotide sequence of the DNA fragment and the flanking ADE1 sequences.

TABLE 3

| PCR amplifications performed to generate pAM493 |   |                                   |                                   |                              |
|---|---|-----------------------------------|-----------------------------------|------------------------------|
| PCR Round                                       | Template  | Primer 1                          | Primer 2                          | PCR Product                  |
| 1   | 100 ng of Y007 genomic DNA  | 61-67-CPK009-G<br>(SEQ ID NO: 31) | 61-67-CPK010-G<br>(SEQ ID NO: 32) | ADE1 <sup>-825 to -226</sup> |
|   | 100 ng of Y002 genomic DNA  | 61-67-CPK011-G<br>(SEQ ID NO: 33) | 61-67-CPK012-G<br>(SEQ ID NO: 34) | ADE1 <sup>-225 to 653</sup>  |
|   |   | 61-67-CPK047-G<br>(SEQ ID NO: 54) | 61-67-CPK064-G<br>(SEQ ID NO: 69) | IDI1                         |
|   |   | 61-67-CPK052-G<br>(SEQ ID NO: 57) | 61-67-CPK065-G<br>(SEQ ID NO: 70) | P <sub>GAL</sub>             |
|   |   | 61-67-CPK031-G<br>(SEQ ID NO: 48) | 61-67-CPK053-G<br>(SEQ ID NO: 58) | tHMGR                        |
| 2   | 100 ng each of ADE1 <sup>-825 to -226</sup> and ADE1 <sup>-225 to 653</sup> purified PCR products | 61-67-CPK009-G<br>(SEQ ID NO: 31) | 61-67-CPK012-G<br>(SEQ ID NO: 34) | ADE1 <sup>-825 to 653</sup>  |
|   | 100 ng each of IDI1 and P <sub>GAL</sub> purified PCR products                                    | 61-67-CPK047-G<br>(SEQ ID NO: 54) | 61-67-CPK052-G<br>(SEQ ID NO: 57) | IDI1-P <sub>GAL</sub>        |
| 3   | 100 ng each of IDI1-P <sub>GAL</sub> and tHMGR purified PCR products                              | 61-67-CPK031-G<br>(SEQ ID NO: 48) | 61-67-CPK047-G<br>(SEQ ID NO: 54) | IDI1-P <sub>GAL</sub> -tHMGR |

(GAL1 nucleotide position -1 to -668) (P<sub>GAL</sub>), and a truncated ORF of the HMG1 gene of *Saccharomyces cerevisiae* (HMG1 nucleotide positions 1586 to 3323) (tHMGR), into the TOPO Zero Blunt II cloning vector. Vector pAM468 was generated by inserting DNA fragment ADE1<sup>-825 to 653</sup>, which comprises a segment of the wild-type ADE1 locus of *Saccharomyces cerevisiae* that extends from nucleotide position -225 to position 653 and harbors a non-native internal XmaI restriction site between bases -226 and -225, into the TOPO TA pCR2.1 cloning vector. DNA fragments IDI1-P<sub>GAL</sub>-tHMGR and ADE1<sup>-825 to 653</sup> were generated by PCR amplification as outlined in Table 3. For the construction of pAM493, 400 ng of pAM473 and 100 ng of pAM468 were digested to completion using XmaI restriction enzyme, DNA fragments corresponding to the IDI1-P<sub>GAL</sub>-tHMGR insert

and the linearized pAM468 vector were gel purified, and 4 molar equivalents of the purified insert was ligated with 1 molar equivalent of the purified linearized vector, yielding vector pAM493. FIG. 7C shows a map of the IDI1-P<sub>GAL</sub>-tHMGR insert, and SEQ ID NO: 11 shows the nucleotide sequence of the DNA fragment and the flanking ADE1 sequences.

otide position 504 to position 1103, a HISMX marker, and a non-native XmaI restriction site between the HIS3<sup>504 to -1103</sup> sequence and the HISMX marker, into the TOPO TA pCR2.1 cloning vector. DNA fragments ERG10-P<sub>GAL</sub>-ERG12 and HIS3<sup>-32 to -1000</sup>-HISMX-HIS3<sup>504 to -1103</sup> were generated by PCR amplification as outlined in Table 4. For construction of pAM495, 400 ng of pAM474 and 100 ng of pAM469 were digested to completion using XmaI restriction enzyme, DNA fragments corresponding to the ERG10-P<sub>GAL</sub>-ERG12 insert and the linearized pAM469 vector were gel purified, and 4 molar equivalents of the purified insert was ligated with 1 molar equivalent of the purified linearized vector, yielding vector pAM495. FIG. 7D shows a map of the ERG10-P<sub>GAL</sub>-ERG12 insert, and SEQ ID NO: 12 shows the nucleotide sequence of the DNA fragment and the flanking HIS3 sequences.

(GAL1 nucleotide position -1 to -668) (P<sub>GAL</sub>), and the ORF of the ERG19 gene of *Saccharomyces cerevisiae* (ERG19 nucleotide position 1 to 1341) (ERG19), into the TOPO Zero Blunt II cloning vector. Vector pAM470 was generated by inserting DNA fragment LEU2<sup>-100 to -450</sup>-HISMX-LEU2<sup>1096 to 1770</sup>, which comprises two segments of the LEU2 locus of *Saccharomyces cerevisiae* that extend from nucleotide position -100 to position 450 and from nucleotide position 1096 to position 1770, a HISMX marker, and a non-native XmaI restriction site between the LEU2<sup>1096 to 1770</sup> sequence and the HISMX marker, into the TOPO TA pCR2.1 cloning vector. DNA fragments ERG8-P<sub>GAL</sub>-ERG19 and LEU2<sup>-100 to -450</sup>-HISMX-LEU2<sup>1096 to 1770</sup> were generated by PCR amplification as outlined in Table 5. For the construction of pAM497, 400 ng of pAM475 and 100 ng of pAM470 were digested to completion using XmaI restriction enzyme, DNA

TABLE 4

| PCR reactions performed to generate pAM495 |  |   |  |   |
|--|--|---|--|---|
| PCR Round                                  | Template   | Primer 1  | Primer 2   | PCR Product   |
| 1  | 100 ng of Y007 genomic DNA   | 61-67-CPK013-G<br>(SEQ ID NO: 35)<br>61-67-CPK017-G<br>(SEQ ID NO: 39)<br>61-67-CPK035-G<br>(SEQ ID NO: 50)<br>61-67-CPK057-G<br>(SEQ ID NO: 62)<br>61-67-CPK040-G<br>(SEQ ID NO: 51) | 61-67-CPK014alt-G<br>(SEQ ID NO: 36)<br>61-67-CPK018-G<br>(SEQ ID NO: 40)<br>61-67-CPK056-G<br>(SEQ ID NO: 61)<br>61-67-CPK058-G<br>(SEQ ID NO: 63)<br>61-67-CPK059-G<br>(SEQ ID NO: 64) | HIS3 <sup>-32 to -1000</sup><br>HIS3 <sup>504 to -1103</sup><br>ERG10<br>P <sub>GAL</sub><br>ERG12<br>HISMX |
|  | 10 ng of plasmid pAM330 DNA **   | 61-67-CPK015alt-G<br>(SEQ ID NO: 37)  | 61-67-CPK016-G<br>(SEQ ID NO: 38)  | HISMX-HIS3 <sup>504 to -1103</sup>  |
| 2  | 100 ng each of HIS3 <sup>504 to -1103</sup> and HISMX PCR purified products                              | 61-67-CPK015alt-G<br>(SEQ ID NO: 37)  | 61-67-CPK018-G<br>(SEQ ID NO: 40)  | ERG10-P <sub>GAL</sub>  |
|  | 100 ng each of ERG10 and P <sub>GAL</sub> purified PCR products  | 61-67-CPK035-G<br>(SEQ ID NO: 50)   | 61-67-CPK058-G<br>(SEQ ID NO: 63)  | HIS3 <sup>-32 to -1000</sup> -HISMX-<br>HIS3 <sup>504 to -1103</sup>  |
| 3  | 100 ng each of HIS3 <sup>-32 to -1000</sup> and HISMX-HIS3 <sup>504 to -1103</sup> purified PCR products | 61-67-CPK013-G<br>(SEQ ID NO: 35)   | 61-67-CPK018-G<br>(SEQ ID NO: 40)  | ERG10-P <sub>GAL</sub> -ERG12   |
|  | 100 ng each of ERG10-P <sub>GAL</sub> and ERG12 purified PCR products                                    | 61-67-CPK035-G<br>(SEQ ID NO: 50)   | 61-67-CPK040-G<br>(SEQ ID NO: 51)  |   |

\*\* The HISMX marker in pAM330 originated from pFA6a-HISMX6-PGAL1 as described by van Dijken et al. ((2000) Enzyme Microb. Technol. 26(9-10): 706-714).

Plasmid pAM497 was generated by inserting the ERG8-P<sub>GAL</sub>-ERG19 insert of pAM475 into vector pAM470. Vector pAM475 was generated by inserting DNA fragment ERG8-P<sub>GAL</sub>-ERG19, which comprises the ORF of the ERG8 gene of *Saccharomyces cerevisiae* (ERG8 nucleotide position 1 to 1512) (ERG8), the genomic locus containing the divergent GAL1 and GAL10 promoter of *Saccharomyces cerevisiae*

fragments corresponding to the ERG8-P<sub>GAL</sub>-ERG19 insert and the linearized pAM470 vector were purified, and 4 molar equivalents of the purified insert was ligated with 1 molar equivalent of the purified linearized vector, yielding vector pAM497. FIG. 7E for a map of the ERG8-P<sub>GAL</sub>-ERG19 insert, and SEQ ID NO: 13 shows the nucleotide sequence of the DNA fragment and the flanking LEU2 sequences.

TABLE 5

| PCR reactions performed to generate pAM497 |                                |                                   |                                   |                              |
|--|--------------------------------|-----------------------------------|-----------------------------------|------------------------------|
| PCR Round                                  | Template                       | Primer 1                          | Primer 2                          | PCR Product                  |
| 1  | 100 ng of Y007 genomic DNA     | 61-67-CPK019-G<br>(SEQ ID NO: 41) | 61-67-CPK020-G<br>(SEQ ID NO: 42) | LEU2 <sup>-100 to 540</sup>  |
|  | 10 ng of plasmid pAM330 DNA ** | 61-67-CPK023-G<br>(SEQ ID NO: 45) | 61-67-CPK024-G<br>(SEQ ID NO: 46) | LEU2 <sup>1096 to 1770</sup> |
|  | 100 ng of Y002 genomic DNA     | 61-67-CPK021-G<br>(SEQ ID NO: 43) | 61-67-CPK022-G<br>(SEQ ID NO: 44) | HISMX                        |
|  |                                | 61-67-CPK041-G<br>(SEQ ID NO: 52) | 61-67-CPK060-G<br>(SEQ ID NO: 65) | ERG8                         |
|  |                                | 61-67-CPK061-G<br>(SEQ ID NO: 65) | 61-67-CPK062-G<br>(SEQ ID NO: 66) | P <sub>GAL</sub>             |

TABLE 5-continued

| PCR reactions performed to generate pAM497 |  |                                   |                                   |                                    |
|--|--|-----------------------------------|-----------------------------------|------------------------------------|
| PCR Round                                  | Template   | Primer 1                          | Primer 2                          | PCR Product                        |
|  |  | (SEQ ID NO: 66)<br>61-67-CPK046-G | (SEQ ID NO: 67)<br>61-67-CPK063-G | ERG19                              |
|  |  | (SEQ ID NO: 53)                   | (SEQ ID NO: 68)                   |                                    |
| 2  | 100 ng each of LEU2 <sup>1096 to 1770</sup> and HISMX purified PCR products                        | 61-67-CPK021-G                    | 61-67-CPK024-G                    | HISMX-LEU2 <sup>1096 to 1770</sup> |
|  | 100 ng each of ERG8 and P <sub>GAL</sub> purified PCR products                                     | (SEQ ID NO: 43)                   | (SEQ ID NO: 46)                   |                                    |
| 3  | 100 ng of LEU2 <sup>-100 to 450</sup> and HISMX-LEU2 <sup>1096 to 1770</sup> purified PCR products | 61-67-CPK041-G                    | 61-67-CPK062-G                    | ERG8-P <sub>GAL</sub>              |
|  | 100 ng each of ERG8-P <sub>GAL</sub> and ERG19 purified PCR products                               | (SEQ ID NO: 52)                   | (SEQ ID NO: 67)                   |                                    |
|  |  | 61-67-CPK019-G                    | 61-67-CPK024-G                    | LEU2 <sup>-100 to 450</sup>        |
|  |  | (SEQ ID NO: 41)                   | (SEQ ID NO: 46)                   | HISMX-LEU2 <sup>1096 to 1770</sup> |
|  |  | 61-67-CPK041-G                    | 61-67-CPK046-G                    | ERG8-P <sub>GAL</sub> -ERG19       |
|  |  | (SEQ ID NO: 52)                   | (SEQ ID NO: 53)                   |                                    |

\*\* The HISMX marker in pAM330 originated from pFA6a-HISMX6-PGAL1 as described by van Dijken et al. ((2000) Enzyme Microb. Technol. 26(9-10): 706-714).

Plasmid pAM584 was generated by inserting DNA fragment GAL7<sup>4 to 1021</sup>-HPH-GAL1<sup>1637 to 2587</sup> into the TOPO ZERO Blunt II cloning vector (Invitrogen, Carlsbad, Calif.). DNA fragment GAL7<sup>4 to 1021</sup>-HPH-GAL1<sup>1637 to 2587</sup> comprises a segment of the ORF of the GAL7 gene of *Saccharomyces cerevisiae* (GAL7 nucleotide positions 4 to 1021) (GAL7<sup>4 to 1021</sup>), the hygromycin resistance cassette (HPH), and a segment of the 3' untranslated region (UTR) of the GAL1 gene of *Saccharomyces cerevisiae* (GAL1 nucleotide positions 1637 to 2587). The DNA fragment was generated by PCR amplification as outlined in Table 6. FIG. 7F shows a map and SEQ ID NO: 102 the nucleotide sequence of DNA fragment GAL7<sup>4 to 1021</sup>-HPH-GAL1<sup>1637 to 2587</sup>.

20 were stitched together in a secondary PCR reaction using 25 ng of each of the gel purified PCR fragments and primers PW287-002-CPK1217 and PW287-002-CPK1232, yielding integration cassette natA-P<sub>CTR3</sub><sup>-1 to -734</sup>.

Additional recombinant integration cassettes were generated by stitching RABits. RABits were generated by inserting 25 DNA fragments of interest (MULEs) into the pMULE Entry vector.

The pMULE Entry vector was PCR amplified using primers K162 (SEQ ID NO: 73) and K163 (SEQ ID NO: 74) and 30 pRYSE Entry vector 8 (SEQ ID NO: 14) as template. The reaction mixture was resolved by gel electrophoresis, and the approximately 2.2 kb vector backbone was gel purified. A

TABLE 6

| PCR reactions performed to generate pAM584 |  |                                     |                                     |  |
|--|--|-------------------------------------|-------------------------------------|--|
| PCR Round                                  | Template   | Primer 1                            | Primer 2                            | PCR Product  |
| 1  | 100 ng of Y002 genomic DNA   | 91-014-CPK236-G<br>(SEQ ID NO: 126) | 91-014-CPK237-G<br>(SEQ ID NO: 127) | GAL7 <sup>4 to 1021</sup><br>GAL1 <sup>1637 to 2587</sup>      |
|  |  | 91-014-CPK232-G<br>(SEQ ID NO: 124) | 91-014-CPK233-G<br>(SEQ ID NO: 125) |  |
|  | 10 ng of plasmid pAM547 DNA **   | 91-014-CPK231-G<br>(SEQ ID NO: 123) | 91-014-CPK238-G<br>(SEQ ID NO: 128) | HPH  |
| 2  | 100 ng each of GAL7 <sup>4 to 1021</sup> and HPH purified PCR products                               | 91-014-CPK231-G<br>(SEQ ID NO: 123) | 91-014-CPK236-G<br>(SEQ ID NO: 126) | GAL7 <sup>4 to 1021</sup> -HPH                                 |
| 3  | 100 ng of each GAL1 <sup>1637 to 2587</sup> and GAL7 <sup>4 to 1021</sup> -HPH purified PCR products | 91-014-CPK233-G<br>(SEQ ID NO: 125) | 91-014-CPK236-G<br>(SEQ ID NO: 126) | GAL7 <sup>4 to 1021</sup> -HPH<br>GAL1 <sup>1637 to 2587</sup> |

\*\* Plasmid pAM547 was generated synthetically, and comprises the HPH cassette, which consists of the coding sequence for the hygromycin B phosphotransferase of *Escherichia coli* flanked by the promoter and terminator of the Tefl gene of *Kluyveromyces lactis*.

Integration cassette natA-P<sub>CTR3</sub><sup>-1 to -734</sup> was generated by 55 PCR amplifying the natA marker using primers PW287-002-CPK1217 (SEQ ID NO: 104) and DE\_PW91-027-CPK262-G (SEQ ID NO: 99) and using plasmid DNA comprising the TEF1 promoter and terminator of *Kluyveromyces lactis* (GenBank accession CR382122 REGIONS: 788874..789380 and 787141..787496, respectively) and the nat resistance marker. In addition, the promoter of the CTR3 gene of *Saccharomyces cerevisiae* was PCR amplified from Y002 genomic DNA from positions -1 to -734 using primers PW287-002-CPK1232 (SEQ ID NO: 100) and DE\_PW91-027-CPK263-G (SEQ ID NO: 101). The 2 PCR products

60 DNA fragment comprising the lacZ coding sequence was generated by digesting to completion pRYSE Entry vector 8 using SphI restriction enzyme, heat inactivating the enzyme (20 min at 65° C.), resolving the reaction mixture by gel electrophoresis, and gel purifying the approximately 0.5 kb DNA fragment. The purified DNA fragment comprising the lacZ coding sequence was ligated with the purified vector backbone, yielding the pMULE Entry vector (see FIG. 8 for a plasmid map).

MULEs were PCR amplified using templates and primers 65 as outlined in Table 7. PCR amplifications were done using the Phusion DNA polymerase (New England Biolabs, Ipswich, Mass.) as per manufacturer's suggested protocol.

TABLE 7

| Amplified MULEs   |  |              |           |
|-------------------|--|--------------|-----------|
| MULE              | Primers  | Template     | Size (bp) |
| 5' ERG20-pGAL1/10 | YD-198-75A (SEQ ID NO: 94)<br>YD-198-75B (SEQ ID NO: 95)   | pAM489       | 1,012     |
| mIS2081           | YD-198-75F (SEQ ID NO: 96)<br>YD-198-75H (SEQ ID NO: 98)   | Genetrix2081 | 1,836     |
| mIS2081T          | YD-198-75G (SEQ ID NO: 97)<br>YD-198-75H (SEQ ID NO: 98)   | Genetrix2081 | 1,668     |
| tTDH3             | RYSE 4 (SEQ ID NO: 77)<br>RYSE 7 (SEQ ID NO: 78)           | RABit63 *    | 311       |
| Hyg               | YD-198-75L (SEQ ID NO: 105)<br>YD-198-75M (SEQ ID NO: 106) | RABit21 *    | 1,962     |
| TRP1              | YD-198-75N (SEQ ID NO: 107)<br>YD-198-75O (SEQ ID NO: 108) | pAM489       | 586       |

\* RABit21 comprises SEQ ID NO: 15, and RABit63 comprises SEQ ID NO: 16.

RABits were generated by inserting the MULEs into the pMULE Entry vector. The PCR reactions were resolved by gel electrophoresis, the MULEs were gel purified, the purified MULEs were treated with T4 polynucleotide kinase (PNK) (New England Biolabs, Ipswich, Mass.) as per manufacturer's suggested protocol, and the PNK was heat inactivated at 65° C. for 20 minutes. The pMULE Entry vector was digested to completion using SphI restriction enzyme, the approximately 2.2 kb pMULE Entry vector backbones (lacking lacZ) was gel purified, the purified pMULE Entry vector backbone was treated with Antarctic Phosphatase (New England Biolabs, Ipswich, Mass.), and the phosphatase was heat inactivated at 65° C. for 20 minutes. The pMULE Entry vector backbone was ligated with each of the amplified MULEs, yielding RABits.

RABits to be stitched (Table 8) were placed together in one tube (333 fmole of each RABit) and digested to completion using LguI restriction enzyme (Fermentas, Glen Burnie, Md.). The restriction enzyme was heat inactivated for 20 minutes at 65° C. The RABit digestion reactions were split into three 30 μL reactions; water, buffer, dNTPs, and DNA polymerase were added to each reaction mixture, and a first round of PCR amplification was initiated. Samples were placed on ice, 0.5 uM of each terminal primer (Table 8) were added to the reaction mixtures, and a second round of PCR amplification was performed. The three PCR reaction mixtures were combined in one tube, the reaction mixtures were resolved by gel electrophoresis, and the PCR products were gel purified. FIG. 7G shows a map of the integration cassettes.

TABLE 8

| PCR Amplification of Integration Cassettes |   |  |
|--|---|--|
| Integration Cassette                       | RABits to be combined                             | Terminal Primers for 2 <sup>nd</sup> Round PCR amplification |
| i00280                                     | 5' ERG20-pGAL1/10 - mIS2081 - tTDH3 - Hyg - TRP1  | YD-198-75A<br>(SEQ ID NO: 94)                                |
| i00281                                     | 5' ERG20-pGAL1/10 - mIS2081T - tTDH3 - Hyg - TRP1 | YD-198-75O<br>(SEQ ID NO: 108)                               |

The first round of PCR amplification was performed as follows: one cycle of denature at 98° C. for 2 minutes; 5 cycles of denature at 98° C. for 30 seconds and anneal/extend at 72° C. for 30 seconds per kilobase PCR product. The second round of PCR amplification was performed as follows: one cycle of denature at 98° C. for 2 minutes; 35 rounds of denature at 98° C. for 12 seconds and anneal/extend at 72° C. for 20 seconds per kilobase PCR product; one cycle of final extend at 72° C. for 7 minutes; and a final hold at 4° C.

### Example 3

This example describes methods for generating *Saccharomyces cerevisiae* strains expressing heterologous isoprene synthases.

*Saccharomyces cerevisiae* strains CEN.PK2-1C (Y002) (MATα; ura3-52; trp1-289; leu2-3,112; his3Δ1; MAL2-8C; SUC2) and CEN.PK2-1D (Y003) (MATalpha; ura3-52; trp1-289; leu2-3,112; his3Δ1; MAL2-8C; SUC2) (van Dijken et al. (2000) Enzyme Microb. Technol. 26(9-10):706-714) were prepared for introduction of inducible MEV pathway genes by replacing the ERG9 promoter with the *Saccharomyces cerevisiae* MET3 promoter, and the ADE1 ORF with the *Candida glabrata* LEU2 gene (CgLEU2). This was done by PCR amplifying the KanMX-P<sub>MET3</sub> region of vector pAM328 (SEQ ID NO: 17) using primers 50-56-pw100-G (SEQ ID NO: 19) and 50-56-pw101-G (SEQ ID NO: 20), which include 45 base pairs of homology to the native ERG9 promoter, transforming 10 μg of the resulting PCR product into exponentially growing Y002 and Y003 cells using 40% w/w Polyethylene Glycol 3350 (Sigma-Aldrich, St. Louis, Mo.), 100 mM Lithium Acetate (Sigma-Aldrich, St. Louis, Mo.), and 10 μg Salmon Sperm DNA (Invitrogen Corp., Carlsbad, Calif.), and incubating the cells at 30° C. for 30 minutes followed by heat shocking them at 42° C. for 30 minutes (Schiestl and Gietz. (1989) *Curr. Genet.* 16, 339-346). Positive recombinants were identified by their ability to grow on rich medium containing 0.5 μg/ml Geneticin (Invitrogen Corp., Carlsbad, Calif.), and selected colonies were confirmed by diagnostic PCR. The resultant clones were given the designation Y93 (MAT A) and Y94 (MAT alpha). The 3.5 kb CgLEU2 genomic locus was then amplified from *Candida glabrata* genomic DNA (ATCC, Manassas, Va.) using primers 61-67-CPK066-G (SEQ ID NO: 71) and 61-67-CPK067-G (SEQ ID NO: 72), which contain 50 base pairs of flanking homology to the ADE1 ORF, and 10 μg of the resulting PCR product were transformed into exponentially growing Y93 and Y94 cells, positive recombinants were selected for growth in the absence of leucine supplementation, and selected clones were confirmed by diagnostic PCR. The resultant clones were given the designation Y176 (MAT A) and Y177 (MAT alpha).

Strain Y188 was generated by digesting 2 μg of pAM491 and pAM495 plasmid DNA to completion using PmeI restriction endonuclease (New England Biolabs, Beverly, Mass.), and introducing the purified DNA inserts into exponentially growing Y176 cells. Positive recombinants were selected for growth on medium lacking uracil and histidine, and integration into the correct genomic locus was confirmed by diagnostic PCR.

Strain Y189 was generated by digesting 2 μg of pAM489 and pAM497 plasmid DNA to completion using PmeI restriction endonuclease, and introducing the purified DNA inserts into exponentially growing Y177 cells. Positive recombinants

35

were selected for by growth on medium lacking tryptophan and histidine, and integration into the correct genomic locus was confirmed by diagnostic PCR.

Strain Y238 was generated by mixing approximately  $1 \times 10^7$  cells from strains Y188 and Y189 on a YPD medium plate for 6 hours at room temperature to allow for mating, plating the mixed cell culture to medium lacking histidine, uracil, and tryptophan to select for growth of diploid cells, digesting 2 µg of pAM493 plasmid DNA to completion using PmeI restriction endonuclease, and introducing the purified DNA insert into the exponentially growing diploid cells. Positive recombinants were selected for by growth on medium lacking adenine, and integration into the correct genomic locus was confirmed by diagnostic PCR.

Strains Y210 (Mat A) and Y211 (MAT alpha) were generated by sporulating strain Y238 in 2% Potassium Acetate and 0.02% Raffinose liquid medium, isolating approximately 200 genetic tetrads using a Singer Instruments MSM300 series micromanipulator (Singer Instrument LTD, Somerset, UK), identifying independent genetic isolates containing the appropriate complement of introduced genetic material by their ability to grow in the absence of adenine, histidine, uracil, and tryptophan, and confirming the integration of all introduced DNA by diagnostic PCR.

Strain Y258 was generated by transforming strain Y211 with pAM404 plasmid DNA. Host cell transformants were selected on synthetic defined media, containing 2% glucose and all amino acids except leucine (SM-glu). Single colonies were transferred to culture vials containing 5 mL of liquid SM-glu lacking leucine, and the cultures were incubated by shaking at 30° C. until growth reached stationary phase. The cells were stored at -80° C. in cryo-vials in 1 mL frozen aliquots made up of 400 uL 50% sterile glycerol and 600 uL liquid culture.

Strains Y225 (MAT A) and Y227 (MAT alpha) were generated by transforming exponentially growing Y210 and Y211 cells, respectively, with 2 µg of pAM426 (SEQ ID NO: 18), which comprises a GAL1 promoter operably linked to the coding sequence of an amorpho-4,11-diene synthase gene that is codon-optimized for expression in *Saccharomyces cerevisiae* (Merke et al. (2000) Ach. Biochem. Biophys. 381: 173-180). Host cell transformants were selected on complete synthetic defined media lacking leucine.

Strain Y337 was generated from strain Y227 by rendering the strain unable to catabolize galactose. To this end, pAM584 plasmid DNA was digested to completion using PmeI restriction endonuclease, and the purified DNA insert  $\text{GAL7}^{4 \rightarrow 1021}\text{-HPH-GAL1}^{1637 \rightarrow 2587}$  was introduced into exponentially growing Y227 cells. Positive recombinants were selected for by growth on solid medium lacking adenine, leucine, lysine, histidine, methionine, uracil, and tryptophan, and containing 900 µg/mL hygromycin B (Sigma, St. Louis, Mo.). Integration into the correct genomic locus was confirmed by diagnostic PCR and by testing the strain for inability to use galactose as a carbon source.

Strain Y615 was generated from strain Y337 by replacing the URA3 open reading frame with the hisG open reading frame from *Salmonella* sp. A DNA fragment comprising the hisG open reading frame was PCR amplified from pAM840 using primers 100-150-KB034-G (SEQ ID NO: 109) and 100-150-KB039-G (SEQ ID NO: 110), and the purified DNA fragment was introduced into exponentially growing Y337 cells. Positive transformants were selected for their ability to grow on medium containing 5-fluoroorotic acid and for their inability to grow on medium lacking uracil. Integration into the correct genomic locus was confirmed by diagnostic PCR.

36

Strain Y1775 was generated from strain Y615 by replacing the kanMX marker with the URA3 marker. A DNA fragment encoding the *S. cerevisiae* URA3 auxotrophic marker was PCR amplified from pAM64 (SEQ ID NO: 103) using primers PW-191-046-CPK1212-G (SEQ ID NO: 111) and PW-191-046-CPK1213-G (SEQ ID NO: 112), and the purified DNA fragment was introduced into exponentially growing Y615 cells. Positive recombinants were selected for by growth on medium lacking adenine, leucine, lysine, histidine, methionine, uracil, and tryptophan. Integration into the correct genomic locus was confirmed by diagnostic PCR.

Strain Y1791 was generated from strain Y1775 by restoring the ERG9 locus. A DNA fragment comprising the ERG9 open reading frame was PCR amplified from Y002 genomic DNA using primers PW-191-015-CPK947-G (SEQ ID NO: 113) and PW-191-015-CPK950-G (SEQ ID NO: 114), and the purified DNA fragment was introduced into exponentially growing Y1775 cells. Positive transformants were selected for their ability to grow on medium containing 5-fluoroorotic acid and for their inability to grow on medium lacking uracil. Integration into the correct genomic locus was confirmed by diagnostic PCR.

Strain Y1856 was generated from strain Y1791 by restoring the GAL1, GAL10, and GAL7 locus. A DNA fragment comprising the GAL1, GAL10, and GAL7 genomic region was PCR amplified from Y002 genomic DNA using primers PW-91-093-CPK453-G (SEQ ID NO: 115) and PW-091-144-CPK689-G (SEQ ID NO: 116), and the purified DNA fragment was introduced into exponentially growing Y1791 cells. Positive recombinants were selected for by growth on medium containing 20 g/L galactose and their inability to grow on medium containing 900 µg/mL hygromycin B.

Strain Y1857 was generated from strain Y1856 by disrupting the  $P_{GAL10}$ -ERG20 locus. A DNA fragment encoding the *S. cerevisiae* URA3 auxotrophic marker was PCR amplified from pAM64 (SEQ ID NO: 130) using primers PW-287-002-CPK1215-G (SEQ ID NO: 117) and PW-287-002-CPK1216-G (SEQ ID NO: 118), and the purified DNA fragment was introduced into exponentially growing Y1856 cells. Positive recombinants were selected for by growth on medium lacking adenine, leucine, lysine, histidine, methionine, uracil, and tryptophan. Integration into the correct genomic locus was confirmed by diagnostic PCR.

Strain Y1855 was generated from strain Y1857 by curing the strain from pAM426. Strain Y1857 was propagated in rich Yeast Extract Peptone Dextrose (YPD) medium containing 0.5% leucine (w/v) for 5 days. Every 12 hours, fresh YPD 0.5% LEU was inoculated to an OD<sub>600</sub> of 0.05 using the previous 12 hour growth. After 5 days, the cells were plated on YPD 0.5% leucine agar medium and incubated at 30° C. for 2 days. Cured cells were identified by their ability to grow on minimal medium containing leucine and their inability to grow on medium lacking leucine.

Strain Y736 was generated from strain Y227 by replacing the URA3 marker with a hisG marker. To this end, the hisG marker was PCR amplified using primers KB34 (SEQ ID NO: 75) and KB39 (SEQ ID NO: 76) and plasmid pAM840 as template. Exponentially growing Y211 cells were transformed with the PCR mixture and were then plated on YPD overnight. Host cell transformants were selected by replica plating cells from the YPD plates onto Complete Synthetic Medium (CSM) solid media lacking methionine and leucine and containing 0.1% 5-FOA and 0.1 mg/ml uracil.

Strain Y737 was generated from strain Y736 by transforming exponentially growing cells with pAM940, and selecting host cell transformants on CSM solid media lacking leucine and uracil.

37

Strain Y846 was generated from strain Y737 by curing the strain of pAM426. To this end, strain Y737 was grown for 3 successive nights in rich media supplemented with 6× the usual concentration of leucine, being diluted each morning 100×. The cells were then plated out to well-separated single colonies on rich media and replica plated onto minimal media lacking leucine. Colonies that grew on rich media but did not grow in the absence of leucine were picked, grown, and tested by PCR to verify that the plasmid was no longer present. One such positive colony was stocked as Y846.

Strain Y1858 was generated from strain Y1895 by heterologous expression of an isoprene synthase. To this end, exponentially growing Y1895 cells were transformed with expression plasmid pAM1547. Host cell transformants were selected on CSM agar lacking methionine and leucine and containing 2% glucose.

Strain Y1859 was generated from strain Y1895 by heterologous expression of an isoprene synthase. To this end, exponentially growing Y1895 cells were transformed with expression plasmid pAM1548. Host cell transformants were selected on CSM agar lacking methionine and leucine and containing 2% glucose.

Strain Y1860 was generated from strain Y1895 by heterologous expression of an isoprene synthase. To this end, exponentially growing Y1895 cells were transformed with expression plasmid pAM1549. Host cell transformants were selected on CSM agar lacking methionine and leucine and containing 2% glucose.

Strain Y1861 was generated from strain Y1895 by heterologous expression of an isoprene synthase. To this end, exponentially growing Y1895 cells were transformed with expression plasmid pAM1550. Host cell transformants were selected on CSM agar lacking methionine and leucine and containing 2% glucose.

Strain Y1713 was generated from strain Y846 by replacing the ERG20 gene with the coding sequence for an isoprene synthase. To this end, exponentially growing Y846 cells were transformed with integration cassette i00280. Host cell transformants were selected on YPD agar containing 2% glucose and 300 µg/mL hygromycin B (A.G. Scientific, San Diego, Calif.).

Strain Y1714 was generated from strain Y846 by replacing the ERG20 gene with a coding sequence for a truncated isoprene synthase. To this end, exponentially growing Y846 cells were transformed with integration cassette i00281. Host cell transformants were selected on YPD agar containing 2% glucose and 300 µg/mL hygromycin B (A.G. Scientific, San Diego, Calif.).

Strain Y1732 was generated from strain Y846 by heterologous expression of an isoprene synthase. To this end, exponentially growing Y846 cells were transformed with expression plasmid pAM1549. Host cell transformants were selected on CSM agar lacking methionine and leucine and containing 2% glucose.

Strain Y1733 was generated from strain Y846 by heterologous expression of a truncated isoprene synthase. To this end,

38

exponentially growing Y846 cells were transformed with expression plasmid pAM1550. Host cell transformants were selected on CSM agar lacking methionine and leucine and containing 2% glucose.

Strain Y1837 was generated by transforming exponentially growing Y1713 cells with expression plasmid pAM1549. Host cell transformants were selected on CSM agar lacking methionine and leucine and containing 2% glucose.

Strain Y1838 was generated by transforming exponentially growing Y1714 cells with expression plasmid pAM1550. Host cell transformants were selected on CSM agar lacking methionine and leucine and containing 2% glucose.

Strain 1907 was generated from strain Y1860 by replacing the ERG20 promoter with the nourseothricin resistance marker (natA) and the copper repressible promoter of the CTR3 gene of *Saccharomyces cerevisiae*. To this end, exponentially growing Y1860 cells were transformed with 200 µg of the integration cassette natA-P<sub>CTR3</sub><sup>-1 to -734</sup>. Host cell transformants were selected for growth on rich YPD medium containing 300 µg/mL nourseothricin (Werner Bio-Agents, Jena, Germany).

#### Example 4

This example describes methods for producing isoprene in *Saccharomyces cerevisiae* host strains.

Single colonies of host cell transformants were transferred to culture vials containing 5 mL of Bird Seed Medium containing 0.25 uM CuSO<sub>4</sub>. The following day, 20 mL of Bird Production Medium containing 1.8% galactose, 0.2% glucose, and 32 uM CuSO<sub>4</sub>, with 4 mL isopropylmyristate, was inoculated with host cell transformant Y1858, Y1859, Y1860, or Y1861 to an OD<sub>600</sub> of 0.05. Similarly, 20 mL of Bird Production Medium containing 1.8% galactose and 0.2% glucose, with 4 mL isopropylmyristate, was inoculated with isolates #3, 6, or 9 of host cell transformant Y1907 to an OD<sub>600</sub> of 0.05. To the Y1907 culture, 0.25 uM CuSO<sub>4</sub>, 50 uM CuSO<sub>4</sub>, or 150 uM CuSO<sub>4</sub> was added. The shake flasks were sealed for anaerobic growth and incubated at 30° C. on a rotary shaker at 200 rpm.

After 72 hours of growth, the cultures were assayed for cell growth. At the same time, 200 µL of isopropylmyristate was removed from each flask and were injected directly on an Agilent 7980 gas chromatograph equipped with a flame ionization detector. To expedite run times, the temperature program and column matrix were modified to achieve optimal resolution and the shortest overall runtime (15.0 min). Each 2 µL sample was split 10:1 and was separated using a Varian fused silica CP-PoraBond U PLOT (25 m×0.32 mm×7 µm; length×width×film thickness) column with hydrogen as the carrier gas. The temperature program for the analysis was as follows: the column was initially held at 100° C. for 1 minute, followed by a temperature gradient of 10° C./min to a temperature of 140° C., followed by a temperature gradient of 40° C./min to a temperature of 250° C., followed by holding the column at 250° C. for 6.5 min. Under these conditions, isoprene elutes at 5.8 minutes.

---

#### SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 128

```

<210> SEQ ID NO 1
<211> LENGTH: 1737
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence

```

-continued

<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 1

```
<210> SEQ ID NO 2
<211> LENGTH: 7348
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polynucleotide
```

<400> SEQUENCE: 2

```
tcggcggttt cggtgtatgac ggtggaaaacc tctgacacat gcaggtcccg gagacggtca 60  
cggcttgtct gtaaaggccat gccccggggccca gacaaggcccg tcggggcccg tcggggatgt 120
```

-continued

-continued

-continued

aaggatctt cacctagatc ctttaaatt aaaaatgaa ttttaaatca atctaaga  
tatatgagta aactgggtct gacagttaacc aatgcttaat cagtgaggca cctatctcg  
cgatctgtct atttcgttca tccatagttg cctgactccc cgtcgtgttag ataactacga  
tacgggaggg cttaccatct ggccccagtg ctgcaatgat accgcgagac ccacgctcac  
cggtccaga tttatcagca ataaaccagc cagccgaaag ggccgagcgc agaagtggc  
ctgcaacttt atccgcctcc atccagttca ttaattgttgc cggaaagct agagtaagta  
gttcgcagt taatagtttgc cgcaacgttgc ttgccattgc tacaggcatc gtgggtc  
gctcgctgtt tggtatggct tcattcagct ccggttccca acgatcaagg cgaggatcat  
gatccccat gttgtgcaaa aaagcggtta gctccttcgg tcctccgatc gttgtc  
gtaagttggc cgcaactttca tcactcatgg ttatggcagc actgcataat tctcttactg  
tcatgccatc cgtaagatgc tttctgtga ctggtgagta ctcaaccaag tcattctgag  
aatagtgtat gcccgcaccc agttgtctt gcccggcgtc aatacggat aataccgc  
cacatagcag aactttaaaa gtgctcatca ttggaaaacg ttcttcggg cgaaaactct  
caaggatctt accgctgttgc agatccagtt cgatgtaaacc cactcggtca cccaaactgt  
cttcagcatc ttttactttc accagcggtt ctgggtgagc aaaaacagga aggcaaaaatg  
ccgaaaaaaaaa gggataaagg gcgcacacgg aatgttgaat actcataactc ttcccttttc  
aatattatttgc aagcattttc cagggttatt gtctcatgag cggatacata tttgaatgt  
tttagaaaaaa taaacaataa ggggtccgc gcacatttc cggaaaatgt ccacctgaac  
gaagcatctg tgcttcattt tgtagaaacaa aaatgcacgc cgagagcgct aattttcaa  
acaaagaatc tgagctgcat ttttacagaa cagaaatgca acgcgaaaagc gctatcc  
caacgaagaa tctgtgttca attttgc aacaaaaatgc caacgcgaga ggcgtaattt  
ttcaaaacaaa gaatctgagc tgcatccatc cagaacacaa atgcaacgc agagcgctat  
tttaccaaca aagaatctat acttccatc ttttgcataaa aatgcatcc cgagagcgct  
attttctaa caaagcatct tagattactt ttttccatc ttgtgcgtc tataatgc  
tctcttgata acttttgca ctgttaggtcc gtaaggta gaagaaggct actttgggt  
ctatccatc ttccatccaaa aaagcctgac tccactcccg cggatccatc attactagcg  
aagctgcggg tgcatccatc caagataaag gcaccccgaa ttatattctt taccatgt  
gattgcgcac acttttgaa cagaaatgttgc tagcgatgat gattccatc tggtc  
attatgttgc gtttccatc ttttgcataaa atataactacg tataatgc  
tcgtattgtt ttgcatttcac tctatgcataa gtttccatc caatccatc ttttgc  
taataactacg tataatgcataa aaaaatgttgc aggtcgatgat gatgtcaag  
gaaagggtggaa tgggttaggtt atataggat atagccacaga gatataatgc  
ttttaggtggaa tggttgcgttcc gcaatccatc ttttgcataaa atataactacg  
tgcgttttgc gtttccatc ttttgcataaa atataactacg tataatgc  
gaagtttccatc ttttgcataaa atataactacg tataatgc  
aaaacgcggc cttccgaaaa tgcaacgcga gctgcgcaca tacagctcac  
gcacccatcatc ctgcgtgttgc cctgtatata tataatgc  
gtttatgttgc ttttgcataaa atataactacg tataatgc  
acccctgttgc atattatccc attccatgcg gggatccatc tgcttc  
tttaggttgc ttttgcataaa atataactacg tataatgc  
ctatatgcgttcc cccatccatc ttttgcataaa atataactacg

-continued

```
atttcccttg atattggatc atactaagaa accattatta tcacgtacatt aacctataaa 7320  
aataaggcgta tcacggggcc ctttcgtc 7348
```

```
<210> SEQ ID NO 3
<211> LENGTH: 952
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polynucleotide
```

<400> SEQUENCE: 3

taatgtttt accatagctc ttccatggct accaatttg tggcttgc taacaaattg  
tcataccaa caccactcc atctaccagg ttcccacaat ctaaaaactt cattactcaa  
aagacctcat tggccaatcc taagccctgg agagtaatct gtgcaacctc atcacagttc  
acccagatca cagaacataa ctcttaggagg tcagcaaatt accagcctaa tttgtggAAC  
ttcgagttct tgcaagtcttt ggagaatgtat ttaaaggctcg aaaagttggaa ggaaaaggct  
acaaaaatttag aagaggaggt caggtgtatg atcaacagg tagatactca gccttgtca  
ttgttagagt tgatagacga tgtccaaagg ttggggttga cctataagtt tgaaaaggac  
atcatcaagg ccttgaaaaa catcgttta ttggacgaga ataagaagaa caagtctgac  
ttacacgcca cagcattatc ttccagggtt ttgagacaac acggattcga ggtatctcag  
gacgtttttg aaagatttaa agacaaggag ggaggtttt caggagaatt aaaaggagat  
gtccaaagggtt tggtgtcatt atacgaggcc tcatatttag gttttgaggg tgagaactta  
ttggaaagaag caaggacctt ttcaatcacc cacttgaaga ataacttaaa ggagggaaa  
aataccaagg tagctgagca agtatcacac gcattggaat taccttatca tcagagattg  
cataggttag aagccagatg gttttggat aagtatgaac ctaaggagcc tcaccaccag  
ttatgtttgg agtttagccaa gttggacttt aatatggtcc aacattaca ccagaaagag  
ttccaggact tgcgtcaatgt gttggacagaa atggggaaaatc gctggcttat tg  
952

```
<210> SEQ ID NO 4
<211> LENGTH: 999
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polynucleotide
```

<400> SEQUENCE: 4

|   |     |
|---|-----|
| taatggatgtt accatagctc ttccaaagag ttgcaggact tgtctagatg gttggacagag | 60  |
| atggggattgg catctaaatt agacttcgtc agagataggt tggatgaaagt atatttctgg | 120 |
| gtttagtta tggccctgtga tcctcaattt ggagagtcga gaaaagctgt aacaaaaatg   | 180 |
| tccggattag tcaccataat tgacgacgtt tatgacgttt acggcacatt ggacgagttt   | 240 |
| caattgttta ctgatgcagt tgaaagggtgg gacgtcaatg ccattaatac cttaccagac  | 300 |
| tacatgaaat tggatgtttt acgtttatac aacacagttt acgacacatc ttactcaatc   | 360 |
| ttgaaggaga agggtcacaa caacttgtct tacttgacta aatcttggag ggagttgtgt   | 420 |
| aaaggcctct tacaagagggc caagtggtca aacaacaaaa taatccctgc tttttcaaaa  | 480 |
| tacttggaga acgcctctgt ctcttcatct ggagtagcat tattggcacc ttcatacttc   | 540 |
| tctgtttggcc agcaacaaga ggacatatct gatcagctt tgagatcatt aaccgatttt   | 600 |

-continued

|  |     |
|--|-----|
| cacgggttgg tacgatcttc atgegtcata ttcagattat gcaatgactt agctacttca  | 660 |
| gctgcagagt tagagagagg agagaccacc aactcaatca tctcttatat gcacgaaaac  | 720 |
| gacggcacct cagaggaaca agcaagagag gagttgagga agttaatcga cgcagagtgg  | 780 |
| aaaaagatga acagggagag ggtctctgat tcaacattgt tgccaaaggc ctcatggag   | 840 |
| atagccgtca atatggccag ggtctcacac tgcacatatac agtacggtga cggtttgggt | 900 |
| aggcctgact acgccaccgaa gaataggatc aagttgttgt tgatcgaccc ttttcttatt | 960 |
| aatcagtta tggatgtctg aggaagagct ggcttatttgc                        | 999 |

<210> SEQ ID NO 5  
<211> LENGTH: 932  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 5

taagtagttg accatagctc ttccatggot acccgagttgt tgcgttgca caggccaata 60  
tctttgaccc acaaattgtt caggaatcca ttgc当地aaag tcataacaagc aacaccattg 120  
accttgaagt taaggtgctc tgttatcaacc gagaacgaaa cattcacaga gacagaaaca 180  
gaaacaagga ggtcagctaa ttacgagccaa aactcatggg attatgacta cttgttgtct 240  
tctgacaccc acgagtctat tgaggtatataa aagacaagg caaagaagtt ggaggctgaa 300  
gtcaggaggg agatcaacaa cggaaaaggca gagttcttgc ctttgc当地a gttgtattgac 360  
aacgtacaga ggttgggatt gggttatagg tttgagtcag acataagaag ggcttggac 420  
aggtttgtat ctccaggtgg attcgacgc当地 gttactaaga cctcatttgc当地 tgcttccgt 480  
ttatctttta ggttatttag acagcatggg ttc当地aagtt cacaggaggc attctcagga 540  
ttcaaaagacc agaacggaaa cttttgc当地 aacttgc当地 aggacataaa agccatcttgc 600  
tctttatacg aagcctcatt ttggccctta gaggggtgaga atattttaga cgaggctaaag 660  
gtcttcgcca tatctcactt gaaggagttg tctgaggaga aaataggaaa ggacttagcc 720  
gaacaagttaa accacggcatt ggaatttacca ttgc当地aaga gaaacttgc当地 gtttgc当地 780  
gtctggctca tcgaggccctaa caggaagaaa gaggatgtc当地 atcaggttt attggagttg 840  
gccccatcttag actacaacat gatccagtc当地 gtcttatcaga gagacttgc当地 agaaacttct 900  
aggqtqgtqqa qaqqqaaqa qctqqcttat tq 932

```
<210> SEQ ID NO 6
<211> LENGTH: 970
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polynucleotide
```

<400> SEQUENCE: 6

|  |     |
|--|-----|
| taagttagttg accatagctc ttccctttag agaaaacttct aggtggtgga gaagagtcgg  | 60  |
| attagccact aaattgcact tcgcttaggga taggttaata gagtcattct attgggctgt   | 120 |
| tggagtagct tttgaaccac aatactcaga ttgttaggaac tcagtagcca agatgttctc   | 180 |
| atccgtcacc ataatcgatg acatctacga cgtatacggaa actttggatg aatttggaaatt | 240 |
| attcaactgat gcagtcgaga gatgggacgt aaatgcattt gatgacttgc ctgattacat   | 300 |
| gaagttgtgc ttcttagctt tgtaacaacac cataaaacgag atcgcatacg acaacttggaa | 360 |

-continued

|  |     |
|--|-----|
| ggacaagggt gaaaatataat tgccttactt aaccaaggcc tgggctgatt tgtgtaacgc   | 420 |
| attcttacag gaagcaaaat gggtgtataa caaatcaaca cctactttcg acgagtattt    | 480 |
| tggtaacgct tggaaagtctt catctggacc tttacaatttgc ttatggctt acttcgcgcgt | 540 |
| cgtacaaaac attaagaaag aggagattga taacttgcaa aagtaccacg atatcatctc    | 600 |
| aagaccatca cacatttca gggttatgtaa cgacttggcc tctgcttcag ctgaaatagc    | 660 |
| tagaggagag actgcaaatt cagtttcatg ttacatgagg accaaggta tatacagaaga    | 720 |
| attagcaacc gaatctgtca tgaatttaat cgacgagacc tggaaagaaga tgaacaagga   | 780 |
| aaagttggga gggttttat tcgcaaaacc ttttgcgaa acagccatca atttagccag      | 840 |
| geagtcacac tgcacatatac acaatggtga tgcccacacc tcacctgacg agttgaccag   | 900 |
| gaaaagagtt ttgtcagttt ttactgaacc tatattacct tttgagaggt gaggaagagc    | 960 |
| tggttatttgc  | 970 |

<210> SEQ ID NO 7  
<211> LENGTH: 932  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

|  |     |
|--|-----|
| <400> SEQUENCE: 7  |     |
| taagtagttg accatagctc ttccatggcc accgagttgt tggcttgca caggcctatc       | 60  |
| tctttgaccc acaagttgtt caggaaccct ttgccaaagg taattcaagc cacaccattt      | 120 |
| accttaaagt tgaggtgctc agtctctacc gagaacgtct ctttaccgaa gacagagact      | 180 |
| gaagcttagga gatcagccaa ctacgagccaa aattcatggg actacgactt cttttatca     | 240 |
| tctgacaccc acgagtcaat tgaggtctac aaagacaagg caaagaaatt ggaggcagag      | 300 |
| gtcaggagag aaatcaacaa cgaaaaggcc gagttctaa ctttggtaa gttgtatcgac       | 360 |
| aatgtacaga gattgggatt gggttacagg ttcgagtctg acattaggag ggcttttagac     | 420 |
| aggtttgtt catcaggtgg atttgacggt gtcacaaaaa catcattaca tgcaactgcc       | 480 |
| ttgtcttca gattattaag gcaacatggaa ttcaagtttcttcttgc                     | 540 |
| ttcaaggacc agaacggtaa cttcttagag aatttggaaagg aggacacccaa agcaatttttgc | 600 |
| tcattgtacg aggcttcatt cttggcattt gaaaggatgaa atatcttggaa tgaagcaaga    | 660 |
| gtattcgcta tctctactt gaaggagttt tctgaagaaa aaatcgaaa agaattggca        | 720 |
| gagcaagtaa atcacgcctt agaattgcct ttacacagaa ggacacagag gttagaagca      | 780 |
| gtctggctta ttgaggctta tagaaagaag gaagatgccaa accaagtatt gttggagtttgc   | 840 |
| gcacattttgg actacaacat gatccagtca gtctaccaga gggacttaag agagacccct     | 900 |
| aggtgggtggaa gaagggaaaga gctggcttat tg                                 | 932 |

<210> SEQ ID NO 8  
<211> LENGTH: 970  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

|  |    |
|--|----|
| <400> SEQUENCE: 8  |    |
| taagtagttg accatagctc ttcccttaag agagacctctt aggtgggtggaa gaagagtagg | 60 |

-continued

---

|   |     |
|---|-----|
| attagctact aaattgcact tcgctaagga cagattaatc gaatcatttt actgggcgtt   | 120 |
| cggagtcgt tttgaacctc agtactctga ttgttagaaac tctgttagcca agatgtttc   | 180 |
| tttcgtcacc atcatcgacg atatctacga cgtctatggt atcttgacg aattagagtt    | 240 |
| gtttaccat gctgtcgaaa ggtggatgt taacgcaatt aatgatttcg ctgactacat     | 300 |
| gaagttgtgc ttttagcct tataacaacac catcaacgag atcgcctatg acaactgaa    | 360 |
| ggacaaggga gagaacattt tgccatactt gaccaaggc tgggctgatt tatgtaacgc    | 420 |
| cttttacaa gaggccaagt gtttatacaa caagtctaca ccaaccttcg acgattattt    | 480 |
| cggaaatgcc tggaaatctt catctggacc tttgcaattt atatggcat acttcgcagt    | 540 |
| tgtccagaac atcaaaaagg aggagatcga aaacttgcag aaataccacg acatcatctc   | 600 |
| aaggccatca catatctca ggttgtgcaa cgatttggca tcagcttcg ctgaaatgc      | 660 |
| aagaggagag acagctaatt cagtctcatg ctacatgagg actaaggaa tctcagaaga    | 720 |
| attggccacc gaatcagtca tgaacttgat tgacgagacc tgcaaaaaga tgaacaagga   | 780 |
| gaagttggga ggtatctttgt ttgc当地acc attcgatcgg accggccataa acttggcttag | 840 |
| gcagtctcac tgc当地acc acaacgggtga cggccacaca tcacctgcacg agttgaccag   | 900 |
| aaaaagggtta ttgtctgtta ttacagaacc aatcttaccc ttgagaggt gaggaagagc   | 960 |
| tggcttattt  | 970 |

<210> SEQ ID NO 9  
 <211> LENGTH: 5050  
 <212> TYPE: DNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

|   |      |
|---|------|
| <400> SEQUENCE: 9   |      |
| gtttaaacta ctattagctg aattgccact gctatcgttt ttagtggcgt tagtgctgc    | 60   |
| attcaagac atggaggcg ttattacgcc ggagctcctc gacagcagat ctgatgactg     | 120  |
| gtcaatatat ttttgcattt aggtctgtt tggaaattata ttttggatg acccatctaa    | 180  |
| tgtactggta tcaccagatt tcatgtcgat ttttaaagcg gctgcttgag tcttagcaat   | 240  |
| agcgtcaccac tctggtaat ccttggagg aaccactgac gaagggttgg acagtgcga     | 300  |
| agaggatctt tcctgctttt aattagtcgc gctgggagca gatgacgagt tggggagct    | 360  |
| ggggggcagga ttgctggccg tcgggggtcc tgaatgggtt cttgggtt ccattcttat    | 420  |
| tctgaaaacg gaagaggagt agggaaatatt actggctgaa aataagtctt gaatgaacgt  | 480  |
| atacgcgtat atttctacca atctctcaac actgagtaat ggtagttata agaaagagac   | 540  |
| cgagtttaggg acagtttagag gcgggtggaga tattcctt ggcattgtctg gcgatgataa | 600  |
| aactttcaa acggcagccc cgatctaaaa gagctgacac cccggagtt tgacaattac     | 660  |
| aacaacagaa ttctttctat atatgcacga acttgcataa tggaaagaaat tatgacgtac  | 720  |
| aaactataaa gtaaaatattt tacgttaacac atgggtgtgt tttcaagaga            | 780  |
| ataccatga cgtatgacta agtttaggt ttaatgcagg tgacggaccc atcttcaaa      | 840  |
| cgatttatat cagtggcgat caaattgtta ggtttttttt gttcagcagg tttccgttt    | 900  |
| tgggtcatat gactttgaac caaatggccg gctgcttaggg cagcacataa ggataattca  | 960  |
| cctgccaaga cggcacaggc aactattttt gctaatttgc gtgcgttgg accaggagcg    | 1020 |
| gttagcatgtg ggccttttac acctaataag tccaacatgg caccttggg ttctagaaca   | 1080 |

-continued

-continued

taccttcaga aacggaaaaat actacataga tatcaccgaa ttgttccatg aggtcacctt 3540  
ccaaacccgaa ttgggccaat tgatggactt aatcactgca cctgaagaca aagtgcactt 3600  
gagtaagttc tccctaaaga agcactcctt catagttact ttcaqaactg cttactattc 3660  
tttctacttg cctgtcgcat tggccatgta cggtgcgggt atcacggatg aaaaggattt 3720  
gaaaacaagcc agagatgtct tgattccatt gggtaatac ttccaaattc aagatgacta 3780  
cttagactgc ttccgttaccc cagaacagat cggttaagatc ggtacagata tccaaagataa 3840  
caaattgtct tgggtaatca acaaggcatt ggaacttgc tccgcagaac aaagaaaagac 3900  
tttagacgaa aattacggta agaaggactc agtgcgacaa gccaaatgca aaaagatttt 3960  
caatgacttg aaaattgaac agctatacca cgaatatgaa gagtctattt ccaaggattt 4020  
gaaggccaaa atttctcagg tcgatgagtc tcgtgggttc aaagctgtat tcttaactgc 4080  
gttcttgaac aaagtttaca agagaagcaa atagaactaa cgctaattcga taaaacattt 4140  
gatattcaaac tagataagga ccatgtataa gaactatata cttccaaatataat aatatagtat 4200  
aagctttaag atagtatctc tcgatctacc gttccacgtg actagtccaa ggattttttt 4260  
taacccggaa tataatgtgtt ctgttcgtt atgacggccat atggcagtag tggaaagat 4320  
tctttattgtt aaaaatagctt gtcaccttac gtacaatctt gatccggagc ttttctttttt 4380  
ttggccgattt agaattcggtt cgaaaaaaaga aaaggagagg gccaadaggg agggcattgg 4440  
tgactattgtt gcacgtgagttt atacgtgtt aagcacacaa aggtagctt gatgtatgtt 4500  
gttattttttt tcacaggttgc ttctgggttca ttgggttggaaag tttggggctt gcagagcaca 4560  
ggggccgcac aatgtgtctt agattccgtt gctgacttgc tgggttattt atgtgtgcac 4620  
aatagaaaga gaacaattgtt cccggttattt gcaaggaaaa tttcaagtct tgtaaaagca 4680  
tataaaaataa gttcaggcac tccgaaatac ttgggttggcg tgggttggaaat tcaacatctt 4740  
gaggatgtttt tggctctgggtt caatgatttac ggcattgata tcgtccaaactt gcatggat 4800  
gagtcgtggc aagaataccca agagttccctt ggtttggccat ttattttttt aactcgatattt 4860  
ccaaaagact gcaacataactt actcagtgca gettcacaga aaccttattt gtttattttttt 4920  
tttgttttttattt cagaagcagg tggggcaggtt gaacttttttggg atttggaaactc gattttctgac 4980  
ttgggttggaa ggcacagagag cccggaaagc ttacatttttta tggacttgc tggactgacg 5040  
ccgttttaaac 5050

```
<210> SEQ ID NO 10
<211> LENGTH: 5488
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polynucleotide
```

<400> SEQUENCE: 10

|  |     |
|--|-----|
| gtttaaacct gctaaattcg agtgaaacac aggaagacca gaaaatcctc atttcatcca    | 60  |
| tattaacaat aatttcaaat gtttatttgc attatttcaa actaggaaag acaagcaacg    | 120 |
| aaacgtttt gaaaattttg agtatttca ataaatttgt agaggactca gatattgaaa      | 180 |
| aaaagctaca gcaattaata cttgataaga agagtattga gaaggggcaac ggttcatcat   | 240 |
| ctcatggatc tgcacatgaa caaacaccag agtcaaacgaa cgttgaaattt gaggctactg  | 300 |
| cgc当地atgaa tgacaatatac gacgatgata acaaaccgaa gttatctgat gttagaaagg   | 360 |
| atttaaatgtat ctaaqaqata qtqatqatat ttccataaata atqtaattct atatatqtta | 420 |

-continued

attaccccttt ttgcgaggca tatttatgtt gaaggataag ttttgaccat ccaaagaagg 480  
taatgtggct gtgggttcag ggtccatacc cgggagttt gacaattaca acaacagaat 540  
tctttctata tatgcacgaa cttgtatat ggaagaatt atgacgtaca aactataaag 600  
taaatatccc acgtaacaca tggtgctgtt gtgcttctt ttcaagagaa tccaatgac 660  
gtatgactaa gtttaggatt taatgcagg t gacggacca tctttcaaac gatttataatc 720  
agtggcgtcc aaattgttag gtttgggg ttcagcagg ttcctgttgtt gggcatatg 780  
actttgaacc aatggccgg ctgtctaggc agcacataag gataattcac ctgcaagac 840  
ggcacaggca actattctt ctaattgacg tgcgttggta ccaggagcg tagcatgtgg 900  
gectcttaca cctaataagt ccaacatggc accttgggtt tctagaacag taccaccacc 960  
gatggtacct acttcgtatgg atggcatgaa tacggaaatt ctcaatcac cgtcacttc 1020  
tttcatcaat gttatacagt tggacttgc gacattttgtt gcaggatctt gtcctaatgc 1080  
caagaaaaca gctgtcaactt aattagctgc atgtgcgtta aatccaccaaa cagacccagc 1140  
cattgcagat ccaaccaaat tcttagcaat gttcaactca accaatgcgg aaacatcaact 1200  
tttaaacact tttctgacaa catcaccagg aatagtagct tctgcgacga cactttacc 1260  
acgacccctcg atccagttga tggcagctgg tttttgtcg gtacagtagt taccagaaac 1320  
ggagacaacc tccatatctt cccagccata ctcttctacc atttgcttta atgagtattc 1380  
gacaccctta gaaatcatat tcatacccat tgcgacca gtagttgttc taaatctcat 1440  
gaagagtaaa tctctgcta gacaagttt aatatgtgc agacgtgcaaa atcttgcgtt 1500  
agagttaaaa gctttttaa ttgcgttttgc tccctttctt gagtctaacc atatcttaca 1560  
ggcaccagat ctttcaaaag ttggaaacg gactactggg cctttgtca taccatcctt 1620  
agttaaaaca gttgttgcac caccgcccggc attgattgcc ttacagccac gcatggcaga 1680  
agctacccaa caaccctctg tagttccat tggatataatgta taagatgtac catcgataac 1740  
caagggccct ataacaccaa cgggcaaaagg catgtacatcataacatttt cacaacaaggc 1800  
gccaatacag cggctgttagt cataatttt atatggtaaa cgatcagatg ctaatacagg 1860  
agcttctgccc aaaattgaaa gagccttcct acgtaccgca accgctctcg tagtatcacc 1920  
taattttttc tccaaagcgt acaaaggtaa cttaccgtga ataaccaagg cagcgaccc 1980  
tttgcgttcc aattgttttgc tatttccact acttaataat gttctaaattt cttctaaagg 2040  
acgtatccatc ttatccaaggc tttcaatatc gccccatca tcttcctcac tagatgtga 2100  
aggtcctgtat gagctcgatt ggcagatgaa taaaacttttgc actttcgatc cagaaatgac 2160  
tgtttttatgg ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 2220  
ggtgacttca gtcttcacca attggctgtc agccattata gtttttttgc ttttttttttgc ttttttttttgc 2280  
aagtatagag gtatattaac aatttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 2340  
aataccaaacc gaaaatgttgc aatgttgc ttttttttttgc ttttttttttgc ttttttttttgc 2400  
atatctgttta atagatcaaa aatcatcgat ttttttttttgc ttttttttttgc ttttttttttgc 2460  
aaaaaaactaa tcgcattttt atccatgttgc ttttttttttgc ttttttttttgc ttttttttttgc 2520  
gtggggccggc gttactgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 2580  
tcttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc ttttttttttgc 2640  
gaagaccagg acgcacggag gagagtcttgc ctttttttttgc ttttttttttgc ttttttttttgc 2700  
tcttaatccgt acttcaatat agcaatgac agttaagcgat ttttttttttgc ttttttttttgc 2760

-continued

gaagggtttt ttaggctaag ataatgggc tctttacatt tccacaacat ataagaaga 2800  
ttagatatgg atatgtatat ggttgttgc ccatgtata tgattattaa acttcttcg 2800  
gtccatccaa aaaaaaagta agaatttttgc aaaattcaat ataaatgaaa ctctcaacta 2940  
aactttgttgc gtgttgttgc aaaggaagac ttaggccca aagcaacaa caattacaca 3000  
atacaaacct gcaaattgact gaactaaaaa aacaaaagac cgctgaacaa aaaaccagac 3060  
ctcaaaatgtt cggttataaa ggtatccaa ttacatccc aactcaatgt gtcaaccaat 3120  
ctgagctaga gaaattgtat ggcgttctc aaggtaata cacaattggc ctggccaaa 3180  
ccaacatgtc tttgtcaat gacagagaag atatctactc gatgtcccta actgtttgt 3240  
ctaagttgtat caagagttac aacatcgaca ccaacaaaat tggttagatta gaagtcggta 3300  
ctgaaactctt gattgacaag tccaaatgtc tcaagtcgtt cttgtatgcaat ttgtttgttgc 3360  
aaaacactga cgtcgaaggtt attgacacgc ttaatgcctt acggatgggaa tggttagagac gccattgttag 3420  
tttgccgtat tattgtccatc tacgataagg gtgccgcaag accaaccgggtt ggtgccgtat 3480  
ctgttgcataat gtggatcggtt cctgtatgtc caattgttgc tgactctgtt agaggttctt 3540  
acatggaaaca cgcctacgtt tttacaagc cagatttac cagcgaatataat ccttacgtcg 3600  
atggtcattt ttcatataact tggttacgtca aggtcttgc tcaagtttac aagagttt 3660  
ccaagaaggctt tatttctaaa ggggtggta gcgatccgc tggttggat gcttgaacg 3720  
ttttgaaata ttgcgtactac aacgttttccat atgttccaaat cttgtaaatttgc tcaacaaaat 3780  
catacggtatc attactataat aacgatttca gagccatcc tcaatttttc ccagaagttt 3840  
acgcccgaattt agctactcgc gattatgacg aatctttaac cgtataagaac attgaaaaaa 3900  
ctttgtttaa tggttgcataag ccattccaca aagagagat tgcccaatctt ttgtattttc 3960  
caacaaacac aggttaacatg tacaccgc tctgttgc tgccttgc tctctattaa 4020  
actatgttgg atctgacgac ttacaaggca agcgtttgg tttattttct tacgggttccg 4080  
gttttagtgc atctctatataat tctgttgc tttgttggta cgtccaaatattatcaagg 4140  
aatttagat tactaacaat ttagccaaga gaatcacca gactccaaatgattacgaag 4200  
ctgccccatcgatc attgagagaa aatgcccatt tgaagaagaa cttcaaacctt caaggttcca 4260  
ttttagcattt gcaaaatgggtt gtttactact tgaccaatcgatc gacaaatggat 4320  
cttacgtatgtt taaaaaataat tcttccccca tctgttgc tctgttgc ac ccccttcata 4380  
aatgctttat tttttggca gcctgtttt ttttagtgc tcttgc tttttttttt 4440  
aatcttatataat ctaggaaaatc tctttatataat atacaatgtatataatataat cccggaaagc 4500  
ttttcaatttctt atctttttttt tttttgttctt tttttttgtt tccgggttctt ttgtttttt 4560  
tttgatttgcg taatctccga gcagaaggaa gaacgaagga aggaggacacag acttagatttgc 4620  
gttatataatc gcatatgttgc tggttgc aatcgaaat tccgggttctt cttttttttt 4680  
ctgcacagaa caaaaacccgtt cagggaaacgc agataaaatca tggttgc aacccaaat 4740  
gaacgtgtctt ctactcatcc tagtccgtt gctgccaagc tttttttttt catgcacggaa 4800  
aagcaaaacaa acttggatgttgc ttcattggat gttcgttacca ccaaggaaat tttttttttt 4860  
gttgcgttgc taggtttttttt tttttttttt cttttttttt cttttttttt 4920  
ttttccatggt agggccacagt taagccgcta aaggcattat cccggaaatgttgc tttttttttt 4980  
cttccatggt agggccacagt taagccgcta aaggcattat cccggaaatgttgc tttttttttt 5040  
cttccatggt agggccacagt taagccgcta aaggcattat cccggaaatgttgc tttttttttt 5100  
gggttataca gaatagcaga atggccacac attacgaatgttgc acacccgggtt ggtggccca 5160

-continued

---

ggtattgtta gcgggttgaa gcaggcggcg gaagaagtaa caaaggaacc tagaggcctt 5220  
 ttgatgttag cagaattgtc atgcaggc tccctagcta ctggagaata tactaagggt 5280  
 actgttgaca ttgcgaagag cgacaaagat tttgttatcg gctttattgc tcaaagagac 5340  
 atgggtggaa gagatgaagg ttacgattgg ttgattatga cacccggtgtt gggtttagat 5400  
 gacaaggggag acgcattggg tcaacagtat agaaccgtgg atgatgtggt ctctacagga 5460  
 tctgacatta ttattgtgg gtttaaac 5488

<210> SEQ ID NO 11  
 <211> LENGTH: 4933  
 <212> TYPE: DNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 11

gtttaaacta ctcagttatata taagttcga attgaaggc gaactttat tcgaagtcgg 60  
 agtcaccaca acacttccgc ccatactctc cgaatcctcg tttcctaaag taagttact 120  
 tccacttgta ggccttattat taatgatatac tgaataatcc tctattaggg ttggatcatt 180  
 cagtagcgcg tgcgattgaa aggagtccat gcccgcacgtc gacgtgattha gcgaaggcgc 240  
 gtaaccatttgc tcatgtcttag cagctataga actaacatcc ttgacaccac ttgcggaaagt 300  
 ctcatcaaca tgctcttccct tattactcat tctcttacca agcagagaat gttatctaaa 360  
 aactacgtgt atttcacccctc tttctcgact tgaacacgtc caactcctta agtactacca 420  
 cagccaggaa agaatggatc cagttctaca cgatagcaaa gcagaaaaca caaccacgt 480  
 acccctgttag aagcttcttt gtttacagca cttgatccat gtagccatac tcgaaatttc 540  
 aactcatctg aaactttcc tgaagggttga aaaagaatgc cataagggtc acccgaagct 600  
 tattcacgcc cgggagttat gacaattaca acaacagaat tctttctata tatgcacgaa 660  
 ctgttaatat ggaagaaatt atgacgtaca aactataaag taaatatttt acgtaacaca 720  
 tggtgctgtt gtgcttcttt ttcaagagaa taccaatgac gtatgactaa gtttaggatt 780  
 taatgcagggt gacggaccca tctttcaaacc gatttataatc agtggcggtcc aaattgttag 840  
 gttttgttgg ttcagcagggt ttctctgtgtt gggcatatgc actttgaacc aaatggccgg 900  
 ctgctagggc agcacataag gataattcac ctgccaagac ggcacaggca actattcttgc 960  
 ctaattgacg tgcgttgta ccaggagcgg tagcatgtgg gccttcttaca cctataaagt 1020  
 ccaacatggc accttgcgtt tctagaacag taccaccacc gatggtacct acttcgatgg 1080  
 atggcatggta tacggaaatt ctcaaataatc cgtccacttc ttcatcaat gttatacagt 1140  
 tggaaacttgc gacattttgtt gcaggatctt gtcctaatgc caagaaaaca gctgtcacta 1200  
 aatttagtgc atgtgcgtt aatccaccaaa cagacccacgc cattgcagat ccaaccaat 1260  
 tcttagcaat gttcaactca accaatgcgg aaacatcaat ttttaacact tttctgacaa 1320  
 catcaccagg aatagtagct tctgcgcacga cacttctacc acgacccatcg atccagttga 1380  
 tggcagctgg tttttgtcg gtacagtagt taccagaaac ggagacaacc tccatcatctt 1440  
 cccagccata ctcttctacc atttgttta atgagtttgc gacaccctta gaaatcatat 1500  
 tcaatccat tgcgtcacca gtatgttgc taaatctcat gaagagtaaa tctctgtcta 1560  
 gacaagtttgc aatatgttgc agacgtgcaatcttgcgtt agagttaaa gctttttaa 1620  
 ttgcgttttgc tcccttttgc gaggcttaacc atatcttaca ggcaccaggat ctttcaag 1680

-continued

ttggggaaacg gactactggg ccttgcata taccatccctt agttaaaaca gttgttgac 1740  
caccgccagc attgattgcc ttacagccac gcatggcaga agctacccaa caaccctctg 1800  
tagttgcat tggttatatga taagatgtac catcgataac caaggggcct ataacaccaa 1860  
cgggcaagg catgtaacct ataacatccc cacaacaacg gccaatatacg cggtcgtagt 1920  
cataattttt atatggtaaa cgatcagatg ctaatacagg agtttctgcc aaaattgaaa 1980  
gaggccttc acgttaccgca accgcetctcg tagtacacc taatttttc tccaaagcgt 2040  
acaaaggtaa cttaccgtga ataaccaagg cagcgacccctc tttgtcttc aattgtttt 2100  
tattttccact acttaataat gcttctaatt cttctaaagg acgtattttc ttatccaagc 2160  
tttcaataatc gcgggaaatca tcttcctcac tagatgtatc aggtcctgt gagctcgatt 2220  
gcgcagatga taaacttttgc acttcgatc cagaaatgtac tgttttattt gttaaaactg 2280  
gtgttagaagg cttttgtaca ggagcagtaa aagacttctt ggtgacttca gttttccacca 2340  
atgggtctgc agccattataa gtttttctc cttgacgtt aagtatagag gtatattaac 2400  
aattttttgt tgatactttt atgacatgg aataagaagg aataacaacc gaaaatgttg 2460  
aaagtatttag taaaagtgtt tatgcagttt ttgcattttt atatctgtt aatgcataaa 2520  
aatcategct tcgctgatatttccatc aaataagggtt aaaaaactaa tcgcattattt 2580  
atcctatggt tgtttaattttt attcgttgc ttgaagggtt gtggggccag gttactgcca 2640  
attttcctc ttcataacca taaaagctatc tattgtatc tctttattttt tcggagcagt 2700  
gcggcgccag gcacatctgc gtttcaggaa cggcggccgtt gaagaccagg acgcacccgg 2760  
gagagtttc cgtcgagggtt ctgtcgcccc ctcggccgtt tctaatccgtt acttcaatatt 2820  
agcaatgagc agttaaaggctt attactgaaa gttccaaaga gaaggtttt ttaggctaaag 2880  
ataatggggc tctttacatt tccacaacat ataagtaaga ttagatatgg atatgtatatt 2940  
gggtggatattt ccattgtatata tgattattaa acttcttgc gtccatccaa aaaaaaaatgt 3000  
agaattttttt aaaaattcaat ataaatgact gcccacaaca atagtgatgcc ccattggcata 3060  
gtatctatgtt acgccaattt agtgcacaaac caaacacccgtt aagacatggg ggaagaggtt 3120  
cctgaaattt ttccattaca acaaagacccat aataccggat ctagtgagac gtccaaatgac 3180  
gaaaggccggag aaacatgtttt ttctggcat gatgaggagc aaattaatgtt aatgtatgg 3240  
aattgtatgg ttttggatgtt ggacgataat gctattgggtt ccgggtacca gaaagttttt 3300  
catttatgg aaaaatattgtt aaaaagggtttt ctacatcgat cattctccgtt ctttttttt 3360  
aatgtaccaag gtgtatattttt tttacaacaa agagccactt gaaaaataac tttccctgtt 3420  
ctttggactt acacatgttgc ctctcatccat ctagtgatggt atgacgatggt agggttgg 3480  
ggtaagcttag acgataagat taagggcgctt attactgccc cggtgagaaaa actagatcat 3540  
gtatggatattt ttccagaaga tgaaaactaag acaaggggta agtttcttattt tttaaacaga 3600  
atccattaca tggcaccaag caatgtacca tgggggtgtt acgttttttga ttacatccat 3660  
ttttataaga tcaacgctaa agaaaaacttgc actgtcaacc caaacgttca tgaagtttaga 3720  
gacttcaat ggggttccacc aatgtatggt aaaaactatgtt ttgtcgatccca aagtttacaag 3780  
tttacgcctt ggttttaatgtt tatttgccat gatgacatccat tcaactgggtt ggagcaattt 3840  
gtatggatattt ctgtatgttca aatgtacccatg ctttttttttata gatgtatccat acaacgcgtt 3900  
aataatataatgtt gtttttttttccatccat aatcataataa acttttttttcat gatgtatccat 3960  
aacgtttcat ttcacccatgttca aatgtatggt aataggccac aaaaatccctt agtataatgtt 4020

-continued

---

|  |      |
|--|------|
| aaactttatt ttctttatcc accgggagt cagtctgact cttgcgagag atgaggatgt   | 4080 |
| aataatacta atctcgaaga tgccatctaa tacatataga catacatata tataatata   | 4140 |
| cattctatacatttaccc agattcttg aggttaagacg gttgggttt atctttgca       | 4200 |
| gttggacta ttaagaacaa tcgaatcata acaattgctt acaaagaata cacatacgaa   | 4260 |
| atattaacga taatgtcaat tacgaagact gaactggacg gtatattgcc attggtgcc   | 4320 |
| agaggtaaag ttagagacat atatgaggta gacgctggta cgttgtgtt tggtgctacg   | 4380 |
| gatcgtatct ctgcataatga cgttattatg gaaaacagca ttcctgaaaa ggggatccta | 4440 |
| ttgaccaaac tgcagagtt ctgggtcaag ttccctgtcca acgtatgtcg taatcattg   | 4500 |
| gtcgacatcg ccccggtaa gactatccc gattatctac ctgcaaaatt gagcgaacca    | 4560 |
| aagtacaaaa cgcaactaga agaccgctct ctattggttc acaaacataa actaattcca  | 4620 |
| ttggaagtaa ttgtcagagg ctacatcacc ggatctgtt ggaaagagta cgtaaaaaca   | 4680 |
| ggtactgtgc atgggttgc acaacctcaa ggacttaaag aatctcaaga gttcccgaa    | 4740 |
| ccaatcttca ccccatcgac caaggctgaa caaggtgaa atgacgaaaa catctccct    | 4800 |
| gcccaggccg ctgagctgg ggggtgaagat ttgtcacgtg gagtggcaga actggctgta  | 4860 |
| aaactgtact ccaagtgc aa agattatgt aaggagaagg gcatcatcat cgcagacact  | 4920 |
| aaattgttta aac   | 4933 |

<210> SEQ ID NO 12  
<211> LENGTH: 6408  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 12

|  |      |
|--|------|
| gtttaaacta ttgtgagggt cagttatcc atccagatatacccgagag gaaacttctt     | 60   |
| agcgtctgtt ttcgttccat aaggcagttc atgaggatata ttttcgttat tgaagccag  | 120  |
| ctcgtgaatg cttatgtctg ctgaaactgtt gtccatgtcg cttaggtacg caatctccac | 180  |
| aggctgcaaa gttttgtct caagagcaat gttattgtgc accccgtaat tggtaacaa    | 240  |
| gtttaatctg tgcttgtcc ccagctctgt cgtaacccatc agttcatcgatctgtcaag    | 300  |
| aaatttacta ggaatagtgc catggatcag caaccgagaa tggcaatttc tactcggtt   | 360  |
| cagcaacgc gcataaacgc tgggtgtcc gtagacatatacgtaaatgatag gattatcatt  | 420  |
| cataagtttc agagcaatgt ctttattctg gaacttggat ttatggctct tttggttaa   | 480  |
| tttcgcctga ttcttgatct ctttagctt ctcgacgtgg gccttttct tgccatatgg    | 540  |
| atccgcgtca cggcgttgc ttccatgt tacgtgacg tatttcctt taaaccacga       | 600  |
| cgctttgtct tcattcaacg tttccatgt tttttctt ctattgtttt gctgtggaa      | 660  |
| aaacttatacg aaagatgacg actttttttt aattctcgat ttaagagctt ggtgacgcgt | 720  |
| aggagtcaact gccaggtatc gtttgaacac ggcattagtc agggaaatc taacacagtc  | 780  |
| ctttcccgca atttttttt tctattactc ttggcctctt ctagtacact ctatatttt    | 840  |
| ttatgcctcg gtaatgattt tcattttttt tttttccacc tagcggatga ctctttttt   | 900  |
| ttcttagcga ttggcattat cacataatga attatacatt atataaaagta atgtgatcc  | 960  |
| ttcgaagaat atactaaagt tttagcttgc tcgtccccgc cgggtcaccc ggccagcgac  | 1020 |
| atggaggccc agaataaccct ctttgacagt cttgacgtgc gtagctcagg ggcacatgt  | 1080 |

-continued

---

|  |      |
|--|------|
| gactgtcgcc cgtacattta gcccatacat ccccatgtat aatcatttgc atccatacat    | 1140 |
| tttgatggcc gcacggcgcg aagcaaaaat tacggctct cgctgcagac ctgcgagcag     | 1200 |
| ggaaacgctc ccctcacaga cgcgttgaat tgtccccacg ccgcgeccct gtagagaaaat   | 1260 |
| ataaaagggtt aggattgcc actgagggttc ttctttcata tacttcttt taaaatcttgc   | 1320 |
| cgttgcata gttctcacat cacatccgaa cataaacaac catggcagaa ccagccaaa      | 1380 |
| aaaagcaaaa acaaactgtt caggagcgc aaggcgttat ctcccgatc actaatgaaa      | 1440 |
| ctaaaattca aatcgctatt tcgctgaatg gtggatat tcaaataaaa gattcgattc      | 1500 |
| tccctgaaaa gaaggatgac gatgttagtt cccaagctac tcagtcacag gtcatcgata    | 1560 |
| tccacacagg tggtggcttt ttggatcata tgatccatgc gttggcaaaa cactctggtt    | 1620 |
| ggtctcttat tggtgaatgt attggtgacc tgcacattga cgatcaccat actaccgaag    | 1680 |
| attgcggtat cgcattaggg caagcggtca aagaagcaat ggggtctgtc cgtgggtgaa    | 1740 |
| aaagattcgg tactgggttc gcaccattgg atgaggcgct atcacgtgcc gtagtcgatt    | 1800 |
| tatcttagtag accatttgct gtaatcgacc ttggattgaa gagagagatg attggtgatt   | 1860 |
| tatccactga aatgattcca cacttttgg aaagtttgc ggaggcggcc agaattactt      | 1920 |
| tgcgttgtga ttgtctgaga ggtttcaacg atcaccacag aagtggagatg gcggtcaagg   | 1980 |
| cttggctgt tgccataaga gaagctattt cttagcaatgg caccaatgac gttccctcaa    | 2040 |
| ccaaagggtgt ttgtgtgtga agtactgaca ataaaaagat tcttggat taaaactgt      | 2100 |
| cattttgtata gttttttat attgttagtg ttctatTTTtta atcaaattgtt agcgtgattt | 2160 |
| atatTTTTTt tgcctcgac atcatctgcc cagatgcgaa gttaaatgcg cagaaagtaa     | 2220 |
| tatcatgcgt caatcgatg tgaatgctgg tcgctataact gctgtcgatt cgataactaac   | 2280 |
| gcgcgcattcc acccgggatg gtctgcttaa atttcattct gtcttcgaaa gctgaattga   | 2340 |
| tactacgaaa aattttttt tggttcttt tctatcttta ttacataaaa cttcatacac      | 2400 |
| agttaaagatt aaaaacaact aataaataat gcctatcgca aatttagctt tgaagtccat   | 2460 |
| ggtaaattcg tggttcctgg caataataga tcgtcaattt gttgtttgt ggttagttta     | 2520 |
| ttttcaata attgaaatac tagggatttg attttaaatg ctttattcaa attttttgcg     | 2580 |
| cttaacaaac agcagccagt cccacccaag tctgtttcaa atgtctcgta actaaaatca    | 2640 |
| tcttgcattt tcttttgaa actgtcaattt tgctcttgag taatgtctt tcgtaacaaa     | 2700 |
| gtcaaagagc aaccggccgc accagcacccg gtaagtttg tggagccat tctcaaatca     | 2760 |
| tcgctcagat tttaataag ttctaatcca ggatgagaaa caccgattga gacaaggcagt    | 2820 |
| ccatgatTTTA ttcttatcaa ttccaaatagt tggtcataca gttcatttatt agtttctaca | 2880 |
| gcctcgatcgat cgggtccctt acatTTTactt aacttagtca tgatctctaa gccttgcagg | 2940 |
| gcacattcac ccatggcatc tagaattggc ttccataactt caggaaattt ctgggtgacc   | 3000 |
| aacacacgaa cgcgagcaac aagatTTTTt gtagacccgg gaaattttagt ataggtagg    | 3060 |
| atcatggaa tggctggaa atcatctaag aacttaaaat tggtttgtt tattgttcca       | 3120 |
| ttatgtgagt ctttttcaaa tagcaggca ttaccataag tggccacagc gttatctatt     | 3180 |
| cctgaagggg taccgtgaat acatTTTca cctatgaagg cccattgatt cactatatgc     | 3240 |
| ttatcgTTTT ctgacagctt ttccaaatca ttagatccta ttaacccccca caagtagggcc  | 3300 |
| atagctaagg cgcgtatac agaaaatagag ggcgttgc gcaacccagc accgtgggt       | 3360 |
| aaagtagact taaaagaaaa cttaaatattc ttggcatggg ggcataaggca aacaaacata  | 3420 |
| tacagggaaac aaaacgctgc atggtagtgg aaggattcgg atagttgagc taacaacgg    | 3480 |

-continued

---

|   |      |
|---|------|
| tccaaaagac taacgagttc ctgagacaag ccatcggtgg cttgttgagc cttggccat    | 3540 |
| tttgggagt ttacttgatc ctgcgtgatc gcattgaat cattgtatgg ccacttatga     | 3600 |
| ttaaagctaa tgcgggaa gtccaattca atagtatctg gtgcagatga ctgcgttatt     | 3660 |
| agcaggtagg ttctcaacgc agacacacta gcagcgacgg caggcttggt gtacacagca   | 3720 |
| gagtgttcac caaaaataat aaccttccc ggtgcagaag ttaagaacgg taatgacatt    | 3780 |
| atagttttt ctcccttgacg ttaaagtata gaggtatatt aacaattttt tggtgatact   | 3840 |
| tttatgacat ttgaataaga agtaatacaa accgaaaatg ttgaaagtat tagttaaagt   | 3900 |
| ggttatgcag cttttgcatt tatatatctg ttaatagatc aaaaatcatc gcttcgtga    | 3960 |
| ttaattaccc cagaataaag gctaaaaaac taatcgcatt attatcctat gggtgttaat   | 4020 |
| ttgattcgtt gatttgaagg tttgtgggc cagggttactg ccaatttttc ctcttcataa   | 4080 |
| ccataaaagc tagtattgtt gaatctttat tgttcggagc agtgcggcgc gaggcacatc   | 4140 |
| tgcgtttcag gaacgcgacc ggtgaagacc aggacgcacg gaggagagtc ttccgtcgga   | 4200 |
| gggctgtcgc cgcgtggcg gcttctaattc cgtaatttcaa tatagcaatg agcagttaag  | 4260 |
| cgttattactg aaagttccaa agagaagggtt tttttaggtt aagataatgg ggcttttac  | 4320 |
| atttccacaa catataagta agattagata tggatatgtt tatgggtgtt ttgccatgtt   | 4380 |
| atatgattat taaacttctt tgcgtccatc caaaaaaaaaa gtaagaattt ttgaaaattc  | 4440 |
| aatataaattt tctcagaacg tttacattgt atcgactgcc agaaccacaa ttggttcatt  | 4500 |
| ccagggttct ctatcctcca agacagcagt ggaattgggt gctgttgctt taaaaggcgc   | 4560 |
| cttggcttaag gttccagaat tggatgcattt caaggattt gacgaaattt ttttggtaa   | 4620 |
| cgttctttct gccaatttgg gccaagctcc ggccagacaa gttgcgtttt ctgcgggtt    | 4680 |
| gagtaatcat atcggttcaa gcacagttaa caaggctgtt gcatccgcta tgaaggcaat   | 4740 |
| cattttgggt gctcaatcca tcaaatgtgg taatgctgtt gttgttgtt ctgggtttt     | 4800 |
| tgaatctatg actaacgcac catactacat gccagcagcc cgtgcgggtt ccaaatttgg   | 4860 |
| ccaaactgtt cttgttgatg gtgtcggaaag agatgggtt aacgatgcgt acgatggct    | 4920 |
| agccatgggt gtacacgcag aaaagtgtgc cctgtattttt gatattacta gagaacaaca  | 4980 |
| agacaatttt gccatcgaaat cctacaaaaa atctcaaaaaa tctcaaaagg aaggtaaatt | 5040 |
| cgacaatgaa attgtacctg ttaccattaa gggatttaga ggttgcgtt atactcaat     | 5100 |
| cacgaaggac gaggaaacctg cttagattaca cttttttttt ttgagatctg caaggactgt | 5160 |
| tttccaaaaa gaaaacggta ctgttactgc cgtaacgcgt tctccaatca acgtgggtc    | 5220 |
| tgcagccgtc atcttggttt ccgaaaaagt ttttggaa aagaatttga agcctttggc     | 5280 |
| tattatcaaa gtttgggtt aggccgcgtca tcaaccagctt gatatttacat gggctccatc | 5340 |
| tcttgcgtt ccaaaggctt tggatgcattt gggatccatc gatgttgcgtt ttttgcgtt   | 5400 |
| ctttgaattt aatgaaggcct tttcggttgcg cgggtttttt aacactaaga ttttgcgtt  | 5460 |
| agacccatct aaggtaatg tataatgggg tgcgtttgtt ctaggttacc cattgggtt     | 5520 |
| ttctgggtgtt agagtggtt ttactactgtt atccatctta cagcaagaag gaggtaagat  | 5580 |
| cgggtgttgc cccatgggtt atgggtgggg tgggtgttcc tctattgtca ttgaaaagat   | 5640 |
| atgattacgt tctgcgtt tctcatgtatc tttttcataaa aatacataaa tatataaattt  | 5700 |
| gttttatgtt taacaggcat aatttaaagt ttttttttgcg attcatcgat tttcaggat   | 5760 |
| tcaaacgcgtt aggtgtgcct tttgacttac ttttccgggg agaggctac agaattaccc   | 5820 |

-continued

---

|  |      |
|--|------|
| tccacgttga ttgtctgcga ggcaagaatg atcatcaccg tagtgagagt gcgttcaagg    | 5880 |
| ctcttgcgggt tgccataaga gaagccacct cgcccaatgg taccaacgt gttccctcca    | 5940 |
| ccaaagggtgt tcttatgtag tgacaccgtat tatttaaagc tgcagcacatc gatatatata | 6000 |
| catgtgtata tatgtatacc tatgaatgtc agtaagtatg tatacgaaca gtatgatact    | 6060 |
| gaagatgaca aggtaatgca tcattctata cgtgtcattc tgaacgaggc gcgcttcct     | 6120 |
| tttttctttt tgctttttct tttttttct cttgaactcg agaaaaaaaaaaa tataaaagag  | 6180 |
| atggaggaac gggaaaaagt tagttgttgt gataggtggc aagtggattt ccgtaaagAAC   | 6240 |
| aacaagaaaa gcatttcata ttatggctga actgagcgaa caagtgcaaa atttaagcat    | 6300 |
| caacgacaac aacgagaatg gttatgtcc tcctcaactt agaggaaaac caagaagtgc     | 6360 |
| cagaaataac agtagcaact acaataacaa caacggcgcc gtttaaac                 | 6408 |

<210> SEQ\_ID NO 13  
<211> LENGTH: 6087  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 13

|   |      |
|---|------|
| gtttaaaactt ttccaaatagg tggtagcaa tcgtcttact ttcttaacttt tcttaccttt   | 60   |
| tacatttcag caatataat atatatattt caaggatata ccattctaattt gtctgcccct    | 120  |
| aagaagatcg tcgtttgcc aggtgaccac gttggtaaag aaatcacagc cgaaggccatt     | 180  |
| aaggttctta aagctatttc tgatgttgcgt tccaatgtca agttcgattt cgaaaatcat    | 240  |
| ttaattgggt gtgctgtat cgatgtaca ggtgttccat ttccagatga ggcgtggaa        | 300  |
| gcctccaaga aggctgatgc cggtttgtta ggtgctgtgg gtggctctaa atggggtacc     | 360  |
| ggtagtgtta gacctaaca aggttacta aaaatccgtt aagaacttca attgtacgcc       | 420  |
| aacttaagac catgtaaactt tgcatccgac tctcttttag acttatctcc aatcaagcca    | 480  |
| caatttgcta aaggactgta ctgcgttgcgt gtcagagaat tagtggggagg tatttacttt   | 540  |
| ggtaagagaa aggaagacgt tttagttgcg tctgtccccgc cgggtcaccc ggcagcgac     | 600  |
| atggaggccc agaataccct ccttgacagt cttgacgtgc gcagctcagg ggcgtatgt      | 660  |
| gactgtcgcc cgtacattta gcccatacat ccccatgtat aatcatttgc atccatacat     | 720  |
| tttgatggcc gcacggcgcc aagcaaaaat tacggctctt cgctgcagac ctgcgagcag     | 780  |
| ggaaacgctc ccctcacaga cgcgttgaat tgcgtccccgc cgcgtccct gttagagaaat    | 840  |
| ataaaagggtt aggatttgc actgagggttc ttctttcata tacttccttt taaaatcttgc   | 900  |
| ctaggataca gttctcacat cacatccgaa cataaacaac catggcagaa ccagccaaa      | 960  |
| aaaagcaaaa acaaaactgtt caggagcgca aggcgtttat ctcccgatc actaatgaaa     | 1020 |
| ctaaaattca aatcgctatt tcgctgaatg gtggatataat tcaaataaaa gattcgattt    | 1080 |
| ttcctgcaaa gaaggatgac gatgttagctt cccaaatgtac tcagtcacag gtcatcgata   | 1140 |
| ttcacacagg tggatgtttt ttggatcata tgcgttgcgtt gttggcaaaa cactctggtt    | 1200 |
| ggtctcttat tggatgtatgtt attggtgacc tgcacattga cgatcaccat actaccgaa    | 1260 |
| attgcgttat cgcatttaggg caagcgttca aagaagcaat gggtgtgtc cgtgggtaa      | 1320 |
| aaagattccgg tactgggttc gcaccattgg atgaggcgct atcagtcgcgtt gtagtgcatt  | 1380 |
| tatcttagtag accatttgcgtt gtaatcgacc ttggattgaa gagagagatg attgggtattt | 1440 |

-continued

---

|   |      |
|---|------|
| tatccactga aatgattcca cacttttgg aaagttcgc ggaggcgcc agaattactt          | 1500 |
| tgcgttgta ttgtctgaga ggttcaacg atcaccacag aagttagtgc gcttcagg           | 1560 |
| cttggctgt tgccataaga gaagctattt ctagcatgg caccaatgac gttccctcaa         | 1620 |
| ccaaagggtgt tttgttgta agtactgaca ataaaaagat tcttggtttc aagaacttgt       | 1680 |
| catttgtata gttttttat attgttagtg ttcttattta atcaaatgtt agcgtgattt        | 1740 |
| atatttttt tcgcctcgac atcatctgcc cagatgegaa gttaaatgcg cagaaagtaa        | 1800 |
| tatcatgcgt caatcgatg tgaatgctgg tcgctatact gctgtcgatt cgatactaac        | 1860 |
| gcccgcattcc acccggttt ctcattcaag tggtaactgc tgtaaaaatt aagatattta       | 1920 |
| taaattgaag ctggtcgtt ccgaccaata ccgtaggaa acgtaaatata gctattgtaa        | 1980 |
| aaaaaggaaa agaaaagaaa agaaaaatgt tacatatcgat attgatctt ttcccttgg        | 2040 |
| agaccagtct ttgcgtcaat caaagattcg tttgttctt gtggcctga accgacttga         | 2100 |
| gttaaaatca ctctggcaac atccctttgc aactcaagat ccaattcagc tgcaatgg         | 2160 |
| tttagatgatt caaattgtatg gttaaagcc tcaagctgcg cagtagtaaa tttcttgc        | 2220 |
| catccaggaa cagagccaaa caatttatag ataaatgc aaagttcga ctcattttca          | 2280 |
| gctaagtatgt acaacacagc atttggaccc gcatcaacg tggatgcac gattttct          | 2340 |
| ccgtaaaact gattaatggt gtggcaccaa ctgatgatac gcttggaaat gtcattcatg       | 2400 |
| tagaatattt gaggaaaga gtcacacat gtggcatgaa aagatgttga atccatcatt         | 2460 |
| gtttccctt caaaggatggc gaaatcttt tcaacaatgg cttaacgcgat gacttcaat        | 2520 |
| ctctttggta cgacatgttc aattcttctt ttaaatatgtt cggagggttc cacggtaat       | 2580 |
| tgcataccct gagtgaaact cacatcctt ttaatatcgat tgacaacttag gacacaagct      | 2640 |
| ttcatctgag gccagtccaga gctgtctgcg atttgcgttgc ccatggaaatc atgaccatct    | 2700 |
| tcagcttttc ccattcccc ggcacgtat ccgcacaaaca acgtatcaca agctgaacca        | 2760 |
| gacccttttc ttgttatattt agatattttt gaaatgttgcgtt gttgttgcgtt gataactta   | 2820 |
| gcaattgcag agaccaatgc agcaaaagcc gcaacggagg aagctaaacc agctgtgt         | 2880 |
| ggaaaggatattttt ttccggagac aatgtggatg ttccattgtt gataatgtggg caatgaggcg | 2940 |
| tccttcgatt ccatttcattt tcttaattgg cgttaggtcgat gcagacaattt ttgagttctt   | 3000 |
| tcattgtcgat tgctgtgtgg ttctccattt aaccacaaatg tgcgtcgatc aaactcgat      | 3060 |
| gcagtagccg cagaggtaaa cgttctgagg tcatcttgcg ataaatgcac tgatatggac       | 3120 |
| gaatttgggg gcaatgttcaaa ctgcgtgtcc cttttccccca aatacttaag ggttgcgtat    | 3180 |
| ttgacgggtt cggtaacggc tgctgtgtaa acggtcattt tagtttttc ttcttgcgtt        | 3240 |
| taaaggatattt aggtatattt acaattttttt gttgatattt ttatgacatt tgaataagaa    | 3300 |
| gtaataaaaaa ccgaaaatgt tgaaatgtt agttaaagtgc gttatgcacg ttttgcattt      | 3360 |
| atatatctgt taatagatca aaaatcatcg ctgcgtgtat taatttccccca agaaaataagg    | 3420 |
| ctaaaaaaact aatcgatattt ttatcctatg gttgttattt tgatcgatg atttgttgcgtt    | 3480 |
| ttgtggggcc aggttactgc caattttcc tcttcataac cataaaagct agtattgttag       | 3540 |
| aatctttattt gttcggagca gtgcggcgcc aggcacatct gcgtttcagg aacgcgaccc      | 3600 |
| gtgaagacca ggacgcacgg aggagatct tccgtcgagg ggctgtcgcc cgctcgccgg        | 3660 |
| cttctaatcc gtacttcaat atagcatgc gcaatgttgcg gtttgcgtt gtttgcgtt         | 3720 |
| gagaaggatattttt tttaggttca agataatggg gctcttaca ttccacaaatataatgtaa     | 3780 |
| gatttagatattt ggtatgtat atgggtgtat tgccatgtaa tatgattttt aaacttctt      | 3840 |

-continued

---

|   |      |
|---|------|
| gctccatcc aaaaaaaaaaag taagaatttt tgaaaattca atataaatgt cagagtttag            | 3900 |
| gccttcagt gccccaggaa aagcgtaact agctggtgaa tatttagttt tagatccgaa              | 3960 |
| atatgaagca tttgtagtcg gattatcgaa aagaatgcat gctgtagccc atcctacgg              | 4020 |
| ttcattgcaa gagtctgata agtttgaatg gcgtgtgaaa agtaaacaat ttAAAGATGG             | 4080 |
| ggagtggctg taccatataa gtcctaaaac tggcttcatt cctgtttcga taggcccgtc             | 4140 |
| taagaaccct ttcattgaaa aagttatcgtaa acgttattt agctacttta agcctaacat            | 4200 |
| ggacgactac tgcaatagaa acttgttctgt tattgtatatt ttctctgtatg atgcctacca          | 4260 |
| ttctcaggag gacagcgtta ccgaacatcg tggcaacaga agattgaggat ttcatcgca             | 4320 |
| cagaattgaa gaagttccca aaacagggtt gggctctcg gcaggtttag tcacagttt               | 4380 |
| aactacagct ttggcctcct ttttgtatc ggacctggaa aataatgttag acaaataatag            | 4440 |
| agaagttatt cataatttat cacaagttgc tcattgtcaa gtcagggtt aaattggaaag             | 4500 |
| cgggtttgtat gtageggcgg cagcatatgg atctatcaga tatagaagat tcccaccgc             | 4560 |
| atataatctt aatttgcag atatttggaa tgctacttac ggcagtaaac tggcgcattt              | 4620 |
| ggtaatgaa gaagactgga atataacgt taaaagtaac catttaccc ttggattaaac               | 4680 |
| tttatggatg ggcgatatta agaatggtc agaaacagta aaactggtcc agaaggtaaa              | 4740 |
| aaatttggat gatcgacata tgccggaaag ctgtttttata tatacagaac tcgatcatgc            | 4800 |
| aaattctaga tttatggatg gactatctaa actagatcgta ttacacgaga ctcatgacga            | 4860 |
| ttacagcgat cagatattt agtctcttgc gaggaaatgac tgcacccatggc aaaagtatcc           | 4920 |
| ttagatcaca gaagtttagag atgcgttgc cacaatttgc cgtttttttt gaaaaataac             | 4980 |
| taaagaatct ggtgccgata tcgaaacctcc cgtacaaact agcttattttt atgattggca           | 5040 |
| gacctaaaaa ggagttcttta cttgttaat acctgggtct ggtggttatg acgcattgc              | 5100 |
| agtgttgc aagcaatgt ttgtatcttgc ggctcaaaacc gctgtatgaca aaagattttc             | 5160 |
| taagggttca tggctggatg taactcaggc tgactgggggtt gtttagaaag aaaaagatcc           | 5220 |
| ggaaacttat ctgtataat aacttaaggt agataatagt ggtccatgtg acatctttat              | 5280 |
| aaatgtgaag ttgttgttgc ccgcgttta catctaaacc ttcatcttcc gatagtactt              | 5340 |
| gaaattgttc ctteggcgg catgataaaa ttcttttaat gggtacaagc tacccggcc               | 5400 |
| cgggaaagat tctctttttt tatgtatattt gtacataaac ttataatgtt aaattcataa            | 5460 |
| tagaaacgc acgaaatttac aaaatggaaat atgttcatag ggttagacgaa actatatacg           | 5520 |
| caatctacat acatttatca agaaggagaa aaaggaggat gtAAAGGAAT acaggtaagc             | 5580 |
| aaattgtatac taatggctca acgtgataag gaaaaagaat tgcactttaa cattaatatt            | 5640 |
| gacaaggagg agggccaccac acaaaaagttt aggtgtaaaca gaaaatcatg aaactatgt           | 5700 |
| tcctaaat tttttttttt tatattggag gatttctctt aaaaaaaaaaa aaatacaaca aataaaaaaac  | 5760 |
| actcaatgac ctgaccat tttttttttt gatggagttt aagtcaatac cttcttgcac cattttccat    | 5820 |
| aatgggtaaa gttccctcaa gaattttact ctgtcagaaa cggccttaac gacgtatcg              | 5880 |
| acctccctttt cagttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt | 5940 |
| tcatccttac ccagccatg taaaacataa gaaggttctt gggaaaggcaga tgcacggct             | 6000 |
| gaacccgagg ataatgcgtat atcccttagt gccatcaataa aagattctcc ttccacgtt            | 6060 |
| gcgaaagaaa cgttaacacg tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt  | 6087 |

-continued

```

<211> LENGTH: 2733
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
polynucleotide

<400> SEQUENCE: 14

gtaaaaacgac ggccagtatt aaccctcaact aaagggaact cgaggctt caaaggcg 60
ccacggtcgt gcggattggc agactccata tgctatgcgg catcagagca gattgtactg 120
agagtgcacc atatgcggtg taaaataccg cacagatgcg taaggagaaa ataccgcac 180
aggcgccatt cgccatttcag gctgcgcaac ttttggaaag ggcatcggt gccccctct 240
tcgctattac gccagctggc gaaaggggga tgtgctgcaa ggcgattaag ttggtaacg 300
ccagggtttt ccaggtaacg acgttgtaaa acgacggcca gtgaattcga gctcggtacc 360
cggggatcct ctagegtcga cctgcaggca tgcaagcttgcgtatcat ggtcatagct 420
gtttcctgtg taaaattgtt atccgctcac aattccacac aacatacgag ccggaagcat 480
aaagtgtaaa gcttgggtg cctaattgagt gagctaactc acattaatttgcgtgc 540
tagcgagtca tccaaagcccc tcagcccccc tagcgctgtt aagagcgagc tcccgctgag 600
caataacttag cgtcatagct gtttctggg tggctcgctt gggcgagcg gtatcagctc 660
actcaaaggc ggtataacgg ttatccacag aatcaggggta acgcgaggaa aagaacatgt 720
gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgcgcgtgc 780
ataggctccg ccccccgtac gacatcaca aaaatcgacg ctcaagtcaag ctggcgaa 840
acccgacagg actataaaga taccagggcgtt tttccctgg aagctccctc gtgcgtctc 900
ctgttccgac cctgcggcgtt acccgatacc tggctcgctt tcccttgcgtt gggcggtgg 960
cgcttctca tagctcagc tggtaggtatc tcaatgggtt gtaggtcggtt cgcttcaagc 1020
tgggctgtgt gacgcaaccc cccgttcagc cccggcgtt cgccttatcc ggtactatc 1080
gtcttggattt caacccggta agacacgact tatcgccact ggcagcagcc actggtaaca 1140
ggatttagcag acgcgaggat gtagggcggtt ctacaggtt ctgttgggg tggcttaact 1200
acggctacac tagaaagaaca gtatggta tctgcgtct gctgaagccaa gttacctcg 1260
aaaaaaaggat tggtagctct tgatccggca aacaaaccac cgctggtagc ggtgggtttt 1320
ttgtttgcaaa gcaacggattt acgcgcaaccc aaaaaggatc tcaagaagat cctttgtatc 1380
tttctacggg gtctgacgctt cagttggaaacg aaaaactcagc ttaaggatt ttggcatga 1440
gattatcaaa aaggatcttcc acctagatcc tttaaatattttt aaaaatgaatg tttaaatcaa 1500
tctaaaggat atatgagtaa ctgggtcgca tgcttaccaa tgcttaatca gtgaggcacc 1560
tatctcagcg atctgtctat ttcgttcatc catagttgcc tgcgttcccg tcgtttagat 1620
aactacgata cggggggctt taccatctgg cccggcgtt gcaatgatacc cggcggaccc 1680
acgctcaccg gctcccgattt ttcgttcccg tttaaatattttt aaaaatgaatg tttaaatcaa 1740
aagtggctt gcaactttat cccgttccat ccaggcttattt aattgttgcgcgtt gggaaagctgt 1800
agtaagtatg tggccgttta atatggcgca acacgttgcgtt gcttgcgtt cggcatcg 1860
gggtgttacgc tggccgttta atatggcgca acacgttgcgtt gcttgcgtt cggcatcg 1920
agttacatgatc tcccccgtt tggccgttta atatggcgca acacgttgcgtt gcttgcgtt cggcatcg 1980
tgtcagaatg aagtggcccg cagttgttacgc tttttttttt atggcgacac tggccgtt 2040
tcttactgtc atgccatccg taagatgtt tttttttttt atggcgacac tggccgtt 2100

```

-continued

---

|   |      |
|---|------|
| attctgagaa tagtgtatgc ggcgaccgag ttgctcttgc cggcgtaaa tacggataaa  | 2160 |
| taccggccca catagcagaa cttaaaaagt gtcatcatt ggaaaaacgtt ctccggcg   | 2220 |
| aaaactctca aggatcttac cgctgtttag atccagttcg atgtaaccca ctctgtcacc | 2280 |
| caactgatct tcagcatctt ttacttcac cagcgttct gggtgagcaa aaacaggaaag  | 2340 |
| gaaaaatgcc gaaaaaaagg gaataaggc gacacggaaa tggtaataac tcataattt   | 2400 |
| ccttttcaa tattattgaa gcatttatca gggttatttg ctcatgagcg gttacatatt  | 2460 |
| tgaatgtatt tagaaaaata aacaatagg ggttccgc acatttcccc gaaaagtgcc    | 2520 |
| acctgacgta taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac | 2580 |
| gaggccctt catctcgcc gtttcggta tgacggtgaa aacctctgac acatgcagct    | 2640 |
| cccgagaca gtcacagctt gtcgttaagc ggatgccggg agcagacaag cccgtcagg   | 2700 |
| cgcgtagcg ggtgttggcg ggtgtcgggg ctg                               | 2733 |

<210> SEQ\_ID NO 15  
 <211> LENGTH: 4170  
 <212> TYPE: DNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

&lt;400&gt; SEQUENCE: 15

|   |      |
|---|------|
| ggtcgttcgg ctgcggcgag cggtatcagc tcactcaaag gggttaatac gggtatccac   | 60   |
| agaatcaggg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa   | 120  |
| ccgtaaaaag gccgegttgc tggegtttt ccataggctc egccccctg acgagcatca     | 180  |
| aaaaaatcga cgctcaagtc agagggtggcg aaacccgaca ggactataaa gataccaggc  | 240  |
| gttccccct ggaagctccc tcgtgcgtc tcctgttccg accctgccc ttacccgata      | 300  |
| cctgtccgccc ttctccctt cgggaagegt ggccgtttt catagtcac gctgttaggt     | 360  |
| tctcagttcg gtgttagtgc ttgcgtccaa gctgggtgt gtgcacgaa ccccggtca      | 420  |
| gcccggccgc tgcgectt ccggtaacta tcgtcttgc tccaaacccgg taagacacga     | 480  |
| cttategcca ctggcagcag ccactggtaa caggattagc agagcgaggt atgtaggcg    | 540  |
| tgctacagag ttcttgaagt ggtggctaa ctacggctac actagaagaa cagtatttgg    | 600  |
| tatctgcgtc ctgctgaagc cagttacett cgaaaaaga gttggtagct cttgatccgg    | 660  |
| caaaacaaacc accgctggta gccgtgggtt ttttgggtc aagcagcaga ttacgcccag   | 720  |
| aaaaaaaaagga tctcaagaag atcctttgat cttttctacg gggctgtacg ctcagtggaa | 780  |
| cgaaaaactca cgtaaggaa ttttggtcat gagattatca aaaaggatct tcacctagat   | 840  |
| ccttttaat taaaatgaa gtttaatc aatctaaagt atatatgagt aaacttggc        | 900  |
| gcatgcttac caatgctta tcagtggagg acctatctca gcatctgtc tatttcgttc     | 960  |
| atccatagtt gcctgactgc ccgtcggtta gataactacg atacggggagg gcttaccatc  | 1020 |
| tggccccagt gctgcaatga taccgcgaga cccacgctca cccggctccag atttacgc    | 1080 |
| aataaaacccg ccagccggaa gggccggcg cagaagtggt ctgcacatt tatccgcctc    | 1140 |
| catccagttt attaattgtt gcccggaaac tagagtaagt agttcgccag ttaatagtt    | 1200 |
| gcgcacacgtt gttgcattt ctacaggcat cgtgggtca cgctcggtt ttggatggc      | 1260 |
| ttcattcagc tccgggtccc aacgatcaag gcgagttaca tgatccccca tgggtgca     | 1320 |
| aaaagcggtt agctccttcg gtcctccgat cggtgtcaga agtaagttgg ccgcagtg     | 1380 |

-continued

---

|   |
|---|
| atcaactcatg gttatggcag cactgcataa ttctcttact gtcatgccc catcgatgcg 1440    |
| cttttctgtg actgggtgagt actcaaccaa gtcattctga gaatagtgtt tgccggcacc 1500   |
| gagttgtct tgcggcggt caatacggga taataccgcg ccacatagca gaactttaaa 1560      |
| agtgcctatc attggaaaac gttcttcggg gcgaaaactc tcaaggatct taccgctgtt 1620    |
| gagatccagt tcgatgtaac ccactcggtc acccaactga tcttcagcat cttttacttt 1680    |
| caccagcggt tctgggtgag caaaaacagg aaggcaaaaat gcccggaaaa agggataaag 1740   |
| ggcgacacgg aaatgttgaa tactcatca ttgcctttt caatattattt gaagcattta 1800     |
| tcagggttat tgtctcatga gcggttatcat atttgaatgt atttagaaaa ataaacaaat 1860   |
| aggggttccg cgcacatttc cccgaaaagt gccacctgac gtctaagaaa ccattattat 1920    |
| catgacatta acctataaaa ataggcgtat cacgaggcccc tttcatctcg cgcgttccg 1980    |
| tgatgacggt gaaaacctct gacacatgca gctccggag acagtcacag cttgtctgtt 2040     |
| agcggtatgcc gggagcagac aagcccgtaa gggcgctca gcccgggttg gcgggtgtcg 2100    |
| gggctggtaa aacgacggcc agtatttaacc ctcactaaag ggaactcgag gctttcagc 2160    |
| tcacacgcgg ccagggggag ctttcgacac tagtaataca catcatcg tcacaaggttc 2220     |
| atcaaagtgt tggacagaca actataccag catggatctc ttgtatcggt tctttctcc 2280     |
| cgcctctctcg caataacaat gaacactggg tcaatcatag cctacacagg tgaacagagt 2340   |
| agcggttata cagggttat acgggtattc ctacggcaaa aattttcat ttctaaaaaa 2400      |
| aaaaagaaaa attttcttt ccaacgctag aaggaaaaga aaaatctaataatttattt 2460       |
| tggtgatttt ctgagagttc ctttttcat atatcgaatt ttgaatataa aaggagatcg 2520     |
| aaaaaatttt tctattcaat ctgtttctg gtttatttg atagttttt tttgttattat 2580      |
| tattatggat tagtactggg ttatatgggt tttctgtat aactttttt tatttttagtt 2640     |
| tgtttaatct tattttgagt tacattatag ttccctaact gcaagagaag taacattaaa 2700    |
| aatgaaaaag cctgaactca ccgcgacgtc tgcgagaag ttctgtatcg aaaagttcga 2760     |
| cagcgtctcc gacctgtatgc agctctcgga gggcgaagaa ttcgtgttt tcagttcga 2820     |
| tgttaggggg cgtggatgtc ttctgcgggt aaatagctgc gcccgtggg tctacaaaga 2880     |
| tcgttatgtt tatcgactttt ttgcattcgcc cggcgtcccc attccggaaag tgcttgacat 2940 |
| tggggaaatc agcgagagcc tgacatttg catctcccg cgtgcacagg gtgtcacgtt 3000      |
| gcaagacctg cctgaaaccg aactgcccgc tttctgcag cgggtcgccg aggccatgga 3060     |
| tgcgatgcgt gcggccgatc ttaccccgac gacgggttc ggcccatcg gaccgcaagg 3120      |
| aatcggtcaa tacactacat ggcgtgattt catatgcgcg attgctgatc cccatgtgtt 3180    |
| tcactggcaa actgtgtatgg acgacaccgt cagtgcgtcc tgcgcccagg ctctcgatga 3240   |
| gtgtatgtt tggcccgagg actgccccga agtccggcac ctgcgtgcacg cggatttcgg 3300    |
| ctccaacaat gtccctgacgg acaatggccg cataacagcg gtcattgact ggagcgaggc 3360   |
| gtatgttccgg gattccaaat acgaggtcgc caacatttc ttctggagcc cgtgggttggc 3420   |
| ttgtatggg cagcagacgc gctacttgcg cggaggatcc cggagcttgc caggatcgcc 3480     |
| gcggctccgg gcgtatgtc tccgcattgg tcttgaccaa ctctatcaga gcttgggttga 3540    |
| cgcaatttc gatgtgcag ctggggcgca gggtcgtgc gacgcaatcg tccgtccgg 3600        |
| agccggact gtcggcgta cacaatgcg ccgcagaagc gggccgtct ggacgatgg 3660         |
| ctgtgttagaa gtactcgccg atagtgaaa cggacgcccc agcactcgcc cgagggcaaa 3720    |
| ggaatagtt taacttgata ctactagatt tttcttttcc atttataaaa tttttggta 3780      |

-continued

|  |      |
|--|------|
| taattgaagc tttagaagta tgaaaaatc ctttttttc attcttgca accaaaataa     | 3840 |
| gaagcttctt ttattcattg aaatgatgaa tataaaccta acaaaaagaaa aagactcgaa | 3900 |
| tatcaaacat taaaaaaaaaaa taaaagaggt tatctgtttt cccatTTAGT tggagttgc | 3960 |
| atTTCTAAT agatagaact ctcattaat gtggatttag ttctctgtt cgTTTTTT       | 4020 |
| tgttttgttc tcactgtatt tacatttcta ttttagtattt agttattcat ataatcttaa | 4080 |
| cttctcgagg agctctcgct cgtccaaacgc cggcggacct tgaagagcga gctcccgctg | 4140 |
| agcaataact agcgctcatag ctgtttcctg                                  | 4170 |

<210> SEQ ID NO 16  
<211> LENGTH: 2519  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 16

|   |      |
|---|------|
| ggtcgttcgg ctgcggcgag cggtatcago tcactcaaag gcgtaatac ggTTATCCAC    | 60   |
| agaatcaggg gataacgcag gaaaagaaat gtgagcaaaa ggccagcaaa aggccaggaa   | 120  |
| ccgtaaaaag gccgcgttgc tggcgTTTT ccataggctc cgccccCTG acgagcatca     | 180  |
| caaaaatcga cgctcaagtc agaggTggcg aaacccgaca ggactataaa gataccaggc   | 240  |
| gtttccccct ggaagctccc tcgtgcgcgc tcctgttccc accctGCCG ttacccgata    | 300  |
| cctgtccgcc ttTCTCCCTT cgggaagcgt ggcgcTTTCT catagctcac gctgttaggt   | 360  |
| tctcagttcg gtgttaggtcg ttgcgtccaa gctggcgtgt gtgcacgaac ccccccgtca  | 420  |
| gccccgaccgc tgcgcTTTAT ccggtaacta tcgtcttgc tccaacccgg taagacacga   | 480  |
| cttatacgcca ctggcagcag ccactggtaa caggattagc agagcgaggt atgttaggcgg | 540  |
| tgcTACAGAG ttcttgaagt ggtggcctaa ctacggctac actagaagaa cagtatttgg   | 600  |
| tatctgcgtc ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgatccgg   | 660  |
| caaacaacc accgctggta cgggtggTTTTTGTTC aagcagcaga ttacgcgcag         | 720  |
| aaaaaaaaaggta tctcaagaag atcctttgat cttttctacgggtctgacg ctcagtggaa  | 780  |
| cggaaactca cgttaaggga ttTGGTcat gagattatca aaaaggatct tcacctagat    | 840  |
| ccttttaaat taaaaatgaa gttttaatc aatctaaagt atatatgagt aaacttggc     | 900  |
| gcatgtttac caatgtttaa tcagtggaggc acctatctca gcatgtgtc tatttcgttc   | 960  |
| atccatagtt gcctgactgc ccgtcggtta gataactacg atacgggagg gcttaccatc   | 1020 |
| tggccccagt gctgtcatca taccgcgaga cccacgtca cccggctccag atttatcagc   | 1080 |
| aataaaaccag ccagccggaa gggccggagcg cagaagtggt cctgcaactt tatccgcctc | 1140 |
| catccagtc attaattgtt gcccggaaagc tagagtaagt agttcgccag ttaatagttt   | 1200 |
| ggcacaacgtt gttgcattt ctacaggcat cgtgggtca cgctcgctgt ttggatggc     | 1260 |
| ttcattcagc tccggTTCC AACGATCAAG GCGAGTTACA TGATCCCCA TGGTGTCAA      | 1320 |
| aaaagcgggtt agctccTTCG GTCCCTCGAT CGTTGTAGA AGTAAGTTGG CCAGTGT      | 1380 |
| atcactcatg gttatggcag cactgcataa ttctcttact gtcgtccat ccgtaaagatg   | 1440 |
| cttttctgtg actgggtgagt actcaacca gtcattctga gaatagtgtt tgccggcacc   | 1500 |
| gagttgtct tgccggcgtt caatacggga taataccgcg ccacatagca gaactttaaa    | 1560 |
| agtgcgtcatac attggaaaac gttctcggg gcgaaaactc tcaaggatct taccgtgtt   | 1620 |

-continued

|  |      |
|--|------|
| gagatccagt tcgatgtAAC ccactcgTGC acccaactGA tcttcAGCAT ctTTTACTTT  | 1680 |
| caccAGCGTT tCTGGGTGAG caaaaACAGG aaggAAACAGG gcccAaaaaa aggGAATAAG | 1740 |
| ggcGACACGG aaATGTTGAA tactcatCA ttgcTTTTT caatATTATT gaAGCATTa     | 1800 |
| tcAGGGTTAT tGTCTCATGA gCGGTTACAT atttGAATGT atttGAAaaa atAAACAAAT  | 1860 |
| aggGGTTCCG CGCACATTc CCCGAAAAGT GCCACCTGAC gtctaAGAAA CCATTATTAT   | 1920 |
| catgacatta acCTATAAAA atAGGCGTAT cacGAGGCC tttcatCTG CGCGTTcGG     | 1980 |
| tGATGACGGT gAAAACCTCT gACACATGCA gCTCCCAGG acAGTCACAG CTTGTCGTa    | 2040 |
| AGCGGATGCC GGGAGCAGAC aAGCCCGTCA gGGCGCGTCa GCGGGTGTG GCGGGTGTcG   | 2100 |
| gggCTGGTAA AACGACGGCC AGTATTAACC CTCACTAAAG ggaACTCGAG gCTTTcACG   | 2160 |
| CTCGTCCAAC GCGGGCGGAC ctaAGCCAAT ATCCCCAAA TTATTAAGAG CGCCTCCATT   | 2220 |
| ATTAACTAAA ATTTCACTCA GCACTCACAA TGTATCAGGT ATCTACTACA GATATTACAT  | 2280 |
| gtggcGAAAAA AGACAAGAAC AATGCAATAG CGCATCAAGA AAAAACACAA AGCTTTCAAT | 2340 |
| caatGAATCG AAAATGTcat TAAAATAGTA TATAAATTGA aactaAGTCA TAAAGCTATA  | 2400 |
| AAAAGAAAAT TTATTTAAAT GCAAGATTa AAGTAAATTc ACTTAATCCC CGCGTGTGTT   | 2460 |
| gCGGGCGTT GAAGAGCGAG CTCCCGCTGA GCAATAACTA gCGTCAAGC TGTTTCTG      | 2519 |

<210> SEQ ID NO 17  
<211> LENGTH: 2121  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide  
<220> FEATURE:  
<221> NAME/KEY: modified\_base  
<222> LOCATION: (13)..(13)  
<223> OTHER INFORMATION: a, c, t, g, unknown or other  
<220> FEATURE:  
<221> NAME/KEY: modified\_base  
<222> LOCATION: (2109)..(2109)  
<223> OTHER INFORMATION: a, c, t, g, unknown or other

<400> SEQUENCE: 17

|   |     |
|---|-----|
| gaattcGCC ttntggatgg cggcgTTAGt atcgaatcga cagcAGTata gCGACCAGCa  | 60  |
| tccacatacG attgacGcat gatattactt tctgcGcact taacttcGca tctggcaga  | 120 |
| tGATGTCGAG GCGAAAAAAA atataaatca cgctAACATT tgattAAAT agAACAACTA  | 180 |
| caatATAAAA AAACATATAcA aATGACAAGT tcttgAAAC aAGAAATCTT ttattGTCAG | 240 |
| tactgattAG AAAAACtCAT CGAGCATCAA atGAAACTGC aATTTATTCA tatcaggatt | 300 |
| atcaataAccA tATTTTGAA AAAGCCGTT ctgtAAATGAA ggAGAAACt ACCCGAGGCA  | 360 |
| gttccatAGG atggcaAGAt CCTGGTATGc gtctgcGATT CGACTCGTC caACATCAAT  | 420 |
| acaACCTATT AATTCCCTt CGTCAAAAT aAGGTTATCA aGTGAGAAAT cAccATGAGT   | 480 |
| gacgactgAA tccGGTgAGA atggcAAAG cttaTGCATT tctttcAGA CTTGTTCAAC   | 540 |
| aggCCAGCCA ttacGCTGt catcaaAAATC actcGcatCA accAAACCGT tattcatcG  | 600 |
| tGATTGCGCC tgAGCGAGAC gaaatacGCG atcGCTGTtA aAGGACAAT tacAAACAGG  | 660 |
| aatcGAATGC aAccGGCGCA ggaACACTGC cAGCGCATCA acaatATTT cacCTGAATC  | 720 |
| aggatATTCT tctaataAcT gGAATGCTGT tttGCCGGGG atcGcAGTGG tgAGTAACCA | 780 |
| tGcatcatCA ggAGTACGGA tAAAATGCTT gatGGTcGGA agAGGcATAA ATTCCGTCAG | 840 |
| ccAGTTAGt CTGACCATCT catctgtAAC atcATTGGCA acGCTACCTT TGCCATGTT   | 900 |

-continued

|  |      |
|--|------|
| cagaaaacaac tctggcgcat cggggttccc atacaatcga tagattgtcg cacctgattg   | 960  |
| cccgacatta tcgcgagccc atttataccatataatca gcatccatgt tgaaatttaa       | 1020 |
| tgcggcctc gaaacgtgag tctttcctt acccatgggtt gtttatgttc ggatgtgatg     | 1080 |
| tgagaactgt atcctagcaa gatttaaaaa ggaagtatat gaaagaagaa cctcagtgac    | 1140 |
| aaatcctaacc ttttatatt tctctacagg ggcgcggcgtt ggggacaatt caacgcgtct   | 1200 |
| gtgagggggag cggtttccctg ctgcgaggc tgcaagcgagg agccgttaatt tttgttcgc  | 1260 |
| gccgtgcggc catcaaaatg tatggatcga aatgattata catggggatg tatggctaa     | 1320 |
| atgtacgggc gacagtccaca tcatgcccctt gagctgcgcga cgtcaagact gtcaaggagg | 1380 |
| gtattctggg cctccatgtc gctggccggg tgacccggcg gggacgaggc aagctaaaca    | 1440 |
| gatctgtatct tgaaactgag taatgtcgtc agaatacccg tcaagataag agtataatgt   | 1500 |
| agagtaatat accaagtatt cagcatattc tcccttttctt ttgtataaat cacggaagg    | 1560 |
| atgatttata agaaaaatga atactattac acttcattta ccaccctctg atcttagattt   | 1620 |
| tccaacgata tgtacgttgtt ggtataaggt gagggggtcc acagatataa catcggtttaa  | 1680 |
| tttagtacta acagagactt ttgtcacaac tacatataag tgtacaataa tagtacagat    | 1740 |
| atgacacact tgttaggcgcac acgogcatcc tacggattgc tgacagaaaa aaaggctacg  | 1800 |
| tgaccagaaaa agtcacgtgtt aattttgtaa ctcaccgcattt ctatggggcgcgtgc      | 1860 |
| acactgcact caacaccata aaccttagca acctccaaag gaaatcaccc tataacaaag    | 1920 |
| ccacagttt acaacttagt ctcttatgaa gttacttacc aatgagaaat agaggcttt      | 1980 |
| tctcgagaaa tatgaatatg gatatatata tatatatata tatatatata tatatatgtat   | 2040 |
| aacctgggttctt ttttttagct tttgtatcttctt agcttgggttctt gtaacagttt      | 2100 |
| tgatatcgna agggcgaatt c  | 2121 |

<210> SEQ ID NO 18  
<211> LENGTH: 8425  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

&lt;400&gt; SEQUENCE: 18

|  |     |
|--|-----|
| tcgcgcgttt cgggtatgac ggtaaaaacc tctgacacat gcagctcccg gagacggta     | 60  |
| cagcttgtct gtaagcggtt gcccggagca gacaagcccg tcagggcgcg tcagcggtt     | 120 |
| ttggcgggtt tcggggctgg cttaactatg cggcatcaga gcagattgtt ctgagatgtc    | 180 |
| accatatacgta ctacgtcgta aggccgtttc tgacagatgtt aaattcttgc gggaaatttc | 240 |
| accattatgg gaaatgttc aagaaggat tgacttaaac tccatcaaattt ggtcaggta     | 300 |
| ttgagtggtt ttatgtttt ttttttagat aaaaatctcc aatataatcaa               | 360 |
| taggaatcgatgtt tttctgttta caccatactt tttgtgttgtt gcccctcc            | 420 |
| ttgtcaatataatgtttaaa gtgcattttt ttttccttacatcgcgttgcgttgc            | 480 |
| aatttgcttta cctgtattcc tttactatcc tcccttttctt ctttcttgcataatgtatgt   | 540 |
| agattgcgtatataatgtttcg tctaccctat gaacatattc cattttgtttaa tttcggtcg  | 600 |
| tttcttattttaa gaaatccatttataatgtttt atgtacaatataatgtttt aagagaatct   | 660 |
| ttttaaagcaaa ggattttctt aacttcttcg ggcacagcat caccgacttc ggtggactt   | 720 |
| ttggaaaccac ctaaatcacc agttctgata cctgcatttcaaaacccctttt aactgcattt  | 780 |

-continued

---

|  |      |
|--|------|
| tcaatggcct taccttcttc aggcaagttc aatgacaatt tcaacatcat tgcagcagac      | 840  |
| aagatagtgg cgatagggtc aaccttattc tttggcaaat ctggagcaga accgtggcat      | 900  |
| ggttcgtaca aaccaaatgc ggtgttcttg tctggcaaag aggccaagga cgcagatggc      | 960  |
| aacaaaccca aggaacctgg gataacggag gttcatcg agatgatatc accaaacatg        | 1020 |
| ttgctggta ttataatacc atttaggtgg gttgggttct taactaggat catggcgca        | 1080 |
| gaatcaatca attgatgtt aaccttcaat gttaggaaatt cgttcttgat ggttccctcc      | 1140 |
| acagttttc tccataatct tgaagaggcc aaaagattag ctttatccaa ggaccaaata       | 1200 |
| ggcaatggtg gctcatgtt tagggccatg aaagcggccaa ttcttgtat tctttgcaact      | 1260 |
| tctggAACGG tgtattgttc actatccaa gcgcacccat caccatcg tcctttctc          | 1320 |
| ttaccaaagt aaatacctcc cactaattct ctgacaacaa cgaagtcaatg accttttagca    | 1380 |
| aattgtggct tgattggaga taagtctaaa agagagtcgg atgcaaagtt acatggctt       | 1440 |
| aagttggcgt acaattgaag ttctttacgg attttagta aaccttggc aggtctaaca        | 1500 |
| ctaccggtagcc accattnn accagccaca gcacctaaca aaacggcatc aaccttcttgc     | 1560 |
| gaggcttcca ggcctcatc tgaaatgggg acacctgttag catcgatagc agcaccacca      | 1620 |
| ataaaatgt ttctgaaatc gaacttgaca ttgaaacgaa catcagaat agcttaaga         | 1680 |
| accttaatgg ttccggctgt gatttcttgc ccaacgtgg cacctggccaa aacgacgatc      | 1740 |
| ttcttagggg cagacattac aatggtataat cttgaaata tatataaaaa aagggcgctt      | 1800 |
| agaccgtcg gccaaacaac caattacttgc ttgagaaata gaggataatt atcctataaa      | 1860 |
| tataacgttt ttgaacacac atgaacaagg aagtacagga caattgattt tgaagagaat      | 1920 |
| gtggattttt atgtattgttggatttca ttttaataa ggcaataataa tttaggtatgt        | 1980 |
| ggatataacta gaagttctcc tgcaccgtcg atatgcgggt tgaaataccg cacagatgc      | 2040 |
| taaggagaaa atacccgtc aggaaattgtt aaacgttaat atttgtttaa aatccgcgtt      | 2100 |
| aaatttttgt taaatcagct catttttaa ccaataggcc gaaatccgcgaa aatccctta      | 2160 |
| taaatcaaaa gaatagaccg agatagggtt gagttgttgc ccagtttgg acaagagtc        | 2220 |
| actattaaag aacgtggact ccaacgtcaa agggcgaaaa accgtctatc agggegatgg      | 2280 |
| cccaactacgt ggaagatccg aggecttagt ttaacgtcg cagaattttc gagttatata      | 2340 |
| actttaaaata cgctgaaccc gaacatagaa atatcgaatg ggaaaaaaaaa actgcataaa    | 2400 |
| ggcattaaaa gaggagcgaa tttttttta ataaaaatct taataatcat taaaagataa       | 2460 |
| ataatagtct atatatacgt atataaataa aaaatattca aaaaataaaa taaactatta      | 2520 |
| ttttagcgta aaggatgggg aaagagaaaa gaaaaaaaaattt gatctatcga tttcaattca   | 2580 |
| attcaattta ttcttttcg gataagaaag caacacctgg caatttcttca ctttccaata      | 2640 |
| attccaaaga agcaccacca ccagtagaga catggagac ccggccatg gtttagataga       | 2700 |
| catagggtaa actagcaatg atttgcatttgc atgcttgc tcatctccca ttctcgtaaa      | 2760 |
| attgtcttta octgcatttttggacccctttaaa aaattggccaa agatataataa cagccataag | 2820 |
| taaagggtttt gggatattct ttgttgcataatc atactcttgc tttatgttgc tccaaacgtc  | 2880 |
| ctccacttcc ttataatca gtgtctgac atattcttcg ttgcattgtt attccattcat       | 2940 |
| gtaaaggatttcc aagagatccg aactatgtttt tctcttgc tccgcatttgc ggttgcgttgc  | 3000 |
| gtcatttaat ctcctacccca gaataccact gtaacggaaat aaaggcgag cagatacagc     | 3060 |
| ccactcaact gattcccttgc tgaaaatatc gtcatttgc agataacagg tagttgttag      | 3120 |

-continued

---

|  |      |
|--|------|
| caagtttgca ccaccagtga taataactac gggatcgtgc tcttcagttg tcggtatgtg      | 3180 |
| tccatttcatta gccccatttcg cttctaccat tagattcctt acgaattctt taacgaactc   | 3240 |
| cttcccacag ttgaaataat cagttctacc ttctttggcc agaaaactcct ccatttcgt      | 3300 |
| gttaggtatcc atgaataatt tgtaaatagg ctcatgtat tccggcaacg tgtctaagca      | 3360 |
| ggtgatcgac catctttcca cggcttcagt gaaaatctt aactcctcg aagttccata        | 3420 |
| tgcgtcatac gtgtcatcaa taagtgttat cacagcaact gccttagtga aaaaaactct      | 3480 |
| agctcttgaa tactggggtt cgtaaccaga acctaaaccc caaaaatagc attcaacgat      | 3540 |
| acgatctctc agacatgggg catttttctt aatatcaaat gccttcacc acttgcatac       | 3600 |
| gtgactcaac tcttccttat gtaggctctg caatagattt aactccagg tagctaactt       | 3660 |
| tagcagagtt ttattatggg agtcttggc ctgatagaag ggtatgtact gggcggcctc       | 3720 |
| gatccttggc aatctcttcc acaatggttg cttaaagct ctctggattt cagtgaataa       | 3780 |
| agcggggttt gtactaaacg cgtccttgcataatcgat agccttgatc ttgtgaatcc         | 3840 |
| caggccatct tcaagaattt ttcgccccgg aactctcatg gacgttagcct cataatcc       | 3900 |
| caacaatcc tcaacatcat tcgctaacga ttgtttaaaa gcaccattct tgccttata        | 3960 |
| gttattaaac acatcacacg tgacatagta tccttggtaa cgcacatcgcc taaaccataa     | 4020 |
| gctagacctg tcgccattcc aattatcacc ataggtctcg taaatacatt gcaatgcatt      | 4080 |
| atcaatttca cgttcaaaat gatacggaaat acctaaacgt tgaatctcg aatcagctt       | 4140 |
| caacaaatattt gcatgttca taggaatatc caatgcttcc ttaacaact gtcttacttc      | 4200 |
| cttcctttaga tcgttacta tttgctccac accctgttca acttggtttct cataaatcaa     | 4260 |
| aaattgtatcg ccccaaataag aagggtggaa atttgcattt ggccttataag gtttcttcc    | 4320 |
| agtcaaggcc attgtttct gcagatccgg ggtttttctt ccttgcgtt aaagtataga        | 4380 |
| ggtatattaa caattttttt ttgatacttt tattacattt gaataagaag taatacaaac      | 4440 |
| cgaaaatgtt gaaagtattt gttaaagtgg ttatgcattt ttgcatttataatgtt           | 4500 |
| aatagatcaa aaatcatcgc ttgcgtgatt aattacccca gaaataaggc taaaaacta       | 4560 |
| atcgcattat catcctatgg ttgtttaattt gattcgttca ttgttgcattt tttttttttt    | 4620 |
| ggttactgcc aatttttctt cttcataacc ataaaagctt gtattgtaga atctttattt      | 4680 |
| ttcggagcag tgcggcgca ggcacatcg cgtttcaggaa acgcgcacccgg tgaagacgag     | 4740 |
| gacgcacggaa ggagagtctt cttcggagg gctgtcaccc gctcggcgcc ttctaatccg      | 4800 |
| tactaagatcg tgcttaattt tggccggcga acgtggcggaa aagggaaaggaa aagaaaggcga | 4860 |
| aaggagcgggg cgcttagggcg ctggcaagtg tagcggtcac gctgcgcgtt accaccacac    | 4920 |
| ccggccgcgtt taatgcggcg ctacaggcg cgtcgcgccta ttgccttccaggatcgca        | 4980 |
| actgttggaa agggegatcg gtgcggccctt cttcgctattt acgcccagctg cattaaatgaa  | 5040 |
| tccggccaaacg cgccggggaga ggcgggttgc gtattggcgtt ctcttcgtt tccctcgatca  | 5100 |
| ctgactcgct ggcgcgttgc gttcggctgc gggagcgggtt atcagctcac tcaaaggcgg     | 5160 |
| taatacgggtt atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc     | 5220 |
| agcaaaaaggcc caggaaccgtt aaaaaggccg cgttgcgttgc gttttccat aggctccgccc  | 5280 |
| cccctgacga gcatcacaaa aatcgacgtt caagtcagag gtggcgaaac ccgacaggac      | 5340 |
| tataaaagata ccaggcggtt cccctggaa gctccctcgat ggcgttcctt gttccgaccc     | 5400 |
| tgcgcgttac cggataacctg tccgccttcc tcccttcggg aagcgtggcg ctttcgtatca    | 5460 |
| gctcacgctg taggtatctc agttcggtgtt aggtcgttgc ctccaaagctg ggctgtgtgc    | 5520 |

-continued

---

acgaacccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca 5580  
 acccggttaag acacgactta tcgcccactgg cagcagccac tggtaacagg attagcagag 5640  
 cgaggtatgt aggcggtgct acagagttct tgaagtggtg gcctaactac ggctacacta 5700  
 gaaggacagt atttggtatac tgccgtctgc tgaagccagt tacccctcgga aaaagagttg 5760  
 gtagctttg atccggcaaa caaaccaccc ctggtagccg tggttttttt gtttgcaga 5820  
 agcagattac gcgcagaaaa aaaggatctc aagaagatcc ttgatctt tctacggggt 5880  
 ctgaegctca gtggAACGAA aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa 5940  
 gatatccac ctagatccctt ttaaattaaa aatgaagttt taaatcaatc taaagtatata 6000  
 atgagtaaac ttggctcgac agttaccaat gcttaatcag tgaggcacct atctcagcga 6060  
 tctgtctatt tcgttcatcc atagttgcct gactccccgt cgtgtagata actacgatac 6120  
 gggagggctt accatctggc cccagtgtcg caatgatacc gcgagaccca cgctcacccg 6180  
 ctccagattt atcagaataa aaccagccag ccggaaaggccg cgagcgcaga agtggctcg 6240  
 caactttatc ogcctccatc cagtcattttt attgttgcgg ggaagctaga gtaagtagtt 6300  
 cgccagttaa tagtttgcgc aacgttggc ccattgtctc aggcatcggt gtgtcacgct 6360  
 cgtcggttgg tatggcttca ttcaagtcgg gttcccaacg atcaaggcga gttacatgat 6420  
 ccccatgtt gtgcaaaaaa gccggtagct cttccggcc tccgatcggt gtcagaagta 6480  
 agttggccgc agtgttatca ctcatggta tggcagcact gcataattct cttactgtca 6540  
 tgccatccgt aagatgctt tctgtactg gtgagactc aaccaagtca ttctgagaat 6600  
 agtgtatgcg gcgaccgagt tgctttgcc cggcgtcaat acgggataat accgegcac 6660  
 atagcagaac tttaaaatgt ctcattttt gaaaacgttc ttccggggcga aaactctcaa 6720  
 ggtatcttacc gctgttggaa tccagttcga tgtaacccac tcgtgcaccc aactgatctt 6780  
 cagcatctt tacttcacc aegcgttctg ggtgagcaaa aacaggaagg caaatggcc 6840  
 caaaaaaggaa aataaggccg acacggaaat gttgaataact catactctc cttttcaat 6900  
 attattgttgcgat cattttatcg ggttattgtc tcatgagccg atacatattt gaatgttattt 6960  
 agaaaaataa acaaataaggg gttccgcgc aattttcccg aaaagtgcgc cctgaacgaa 7020  
 gcatctgtgc ttcattttgtt agaacaaaaa tgcaacgcga gagcgtat ttttcaaaaca 7080  
 aagaatctga gtcgttattt tacagaacag aaatgcaacg cgaaagcgtt attttaccaa 7140  
 cgaagaatct gtgttccatt tttgtaaaac aaaaatgcga cgcgagagcg ctaattttc 7200  
 aaacaaagaa tctgagctgc attttacag aacagaaatg caacgcgaga ggcgtatattt 7260  
 accaacaag aatctatact tctttttgt tctacaaaaa tgcattccgc gagcgtatattt 7320  
 ttctcaacaa aacatcttag attactttt ttctcccttg tgcgtctat aatgcgtct 7380  
 cttgataact ttttgcactg taggtccgtt aaggtagaa gaaggctact ttgggtctta 7440  
 ttttcttccataaaaaaa gcctgactcc acttcccgcg tttactgatt actagcgaag 7500  
 ctgcgggtgc atttttcaaa gataaaggca tccccgatta tattctatac cgatgtggat 7560  
 tgcgcataact ttgtgaacag aaagtgtatcg cggtgtatgt tottcatgg tcagaaaatt 7620  
 atgaacgggtt tcttctattt tgcgtctata tactacgtat aggaaatgtt tacatttcg 7680  
 tattgttttc gattcactct atgaatagtt cttactacaa ttttttgc taaagagtaa 7740  
 tactagagat aaacataaaaa aatgttaggg tcgagtttag atgcgttca aaggagcga 7800  
 aggtggatgg gtaggttata tagggatata gcacagagat atatagcaaa gagatactt 7860

-continued

|   |      |
|---|------|
| ttagcaatgt ttgttggaaagc ggtattcgca atatTTTGTG agctcggtac agtccgggtgc  | 7920 |
| gtttttgggt ttttggaaagt gcgtcttcag agcgctttg gttttcaaaa gcgcctgaa      | 7980 |
| gttcctatac tttcttagaga ataggaaacctt cggaatagga acttcaaaagc gtttccgaaa | 8040 |
| acgagcgctt ccgaaaaatgc aacgcgagct gcgcacatac agctcactgt tcacgctcgca   | 8100 |
| cctatatctg cgtgttgctt gtatatatat atacatgaga agaacggcat agtgcggtt      | 8160 |
| tatgcctaaa tgcgtactta tatgcgtcta ttatgttagg atgaaaggta gtctagtacc     | 8220 |
| tcctgtgata ttatcccatt ccatgcgggg ttcgtatgc ttcccttcagc actacccttt     | 8280 |
| agctgttctatatgcgtccca ctccctcaatt ggatttagtct catccttcaa tgctatcatt   | 8340 |
| tcctttgata ttggatcata ctaagaaacc attattatca tgacattaac ctataaaaat     | 8400 |
| aqgcqtagatca cqaqqqcccc tcgqtc  | 8425 |

```
<210> SEQ ID NO 19
<211> LENGTH: 66
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
```

```
<400> SEQUENCE: 19  
gagtgaacct gctgcctggc gtgctctgac tcagttacatt tcatagtgga tggccggcggtt 60  
actatc 66
```

```
<210> SEQ ID NO 20
<211> LENGTH: 65
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      primer
```

```
<400> SEQUENCE: 20  
  
cgtgtatacg tttccgctt ctgtcttcgt tcttttctct tcttccgata tcacaactgt      60  
  
tacqa                               65
```

```
<210> SEQ ID NO 21
<211> LENGTH: 32
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      primer
```

<400> SEQUENCE: 21  
ttttccggat ccataatccc ttttggatttt tg 32

<210> SEQ ID NO 22  
<211> LENGTH: 32  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
primer

<400> SEQUENCE: 22

<210> SEQ ID NO 23  
211 LENGTH 22

-continued

<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 23

gtttaaaacta ctattagctg aattgccact

30

<210> SEQ ID NO 24  
<211> LENGTH: 46  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 24

actgc当地 agacatata cccgggtgtc agctctttt gatcg

46

<210> SEQ ID NO 25  
<211> LENGTH: 46  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 25

ccgatctaaa agagctgaca cccgggatat atgtgtactt tgcagt

46

<210> SEQ ID NO 26  
<211> LENGTH: 30  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 26

gtttaaacgg cgtcagtcca ccagctaaca

30

<210> SEQ ID NO 27  
<211> LENGTH: 30  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 27

gtttaaaactt gctaaattcg agtgaaacac

30

<210> SEQ ID NO 28  
<211> LENGTH: 46  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 28

aaagatgaat tgaaaagctt cccgggtatg gaccctgaaa ccacag

46

<210> SEQ ID NO 29  
<211> LENGTH: 46  
<212> TYPE: DNA

-continued

<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 29

ctgtggtttc agggccata cccgggaagc tttcaattc atcttt 46

<210> SEQ ID NO 30  
<211> LENGTH: 30  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 30

gtttaaaccc aacaataata atgtcagatc 30

<210> SEQ ID NO 31  
<211> LENGTH: 30  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 31

gtttaacta ctcagtatata taagtttcga 30

<210> SEQ ID NO 32  
<211> LENGTH: 70  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 32

atctctcgca agagtcagac tgactcccg gcgtgaataa gttcggttgc acccttatgg 60

catttttttt 70

<210> SEQ ID NO 33  
<211> LENGTH: 70  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 33

aaaaagaatg ccataagggt cacccgaagc ttattcacgc ccgggagtca gtctgactct 60

tgcgagagat 70

<210> SEQ ID NO 34  
<211> LENGTH: 30  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 34

gtttaaacaa tttagtgtct gcgtatgtga 30

-continued

<210> SEQ ID NO 35  
<211> LENGTH: 30  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 35

gtttaaacta ttgtgagggt cagttatttc

30

<210> SEQ ID NO 36  
<211> LENGTH: 44  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 36

ggggggacga ggcaagctaa acttttagtat attcttcgaa gaaa

44

<210> SEQ ID NO 37  
<211> LENGTH: 44  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 37

tttcttcgaa gaatatacta aagtttagct tgcctcgtcc ccgc

44

<210> SEQ ID NO 38  
<211> LENGTH: 60  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 38

caatcaacgt ggagggtaat tctgctagecc tctcccggtt ggtatggcggc gttatgtatcg 60

<210> SEQ ID NO 39  
<211> LENGTH: 60  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 39

cgatactaac gccgcacatcc acccgggaga ggctagcaga attaccctcc acgttgattg 60

<210> SEQ ID NO 40  
<211> LENGTH: 30  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 40

gtttaaacgcg cggcggttgtt gttattgttag

30

&lt;210&gt; SEQ ID NO 41

-continued

<211> LENGTH: 30  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 41

gtttaaacctt ttccaatagg tggtagcaa

30

<210> SEQ ID NO 42  
<211> LENGTH: 55  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 42

gggtgaccccg gcggggacga ggcaagctaa acgtcttcct ttctcttacc aaagt

55

<210> SEQ ID NO 43  
<211> LENGTH: 55  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 43

acttttgtaa gagaaaaggaa gacgtttagc ttgcctcgtc cccgcgggt caccc

55

<210> SEQ ID NO 44  
<211> LENGTH: 62  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 44

aatatcataa aaaaagagaa tcttcccg gtggatggcg gcgttagtat cgaatcgaca

60

gc

62

<210> SEQ ID NO 45  
<211> LENGTH: 62  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 45

gctgtcgatt cgatactaac gccgccatcc acccgaaa gattctcttt ttttatgata

60

tt

62

<210> SEQ ID NO 46  
<211> LENGTH: 45  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 46

gtttaaacgt gttaacgtt ctccgccta cgtgaaaggaa gaatc

45

-continued

<210> SEQ ID NO 47  
<211> LENGTH: 35  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 47

tccccccggg taaaaaaaaa tccttgact agtca

35

<210> SEQ ID NO 48  
<211> LENGTH: 35  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 48

tccccccggg agttatgaca attacaacaa cagaa

35

<210> SEQ ID NO 49  
<211> LENGTH: 30  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 49

tccccccggg tatatatata tcattgttat

30

<210> SEQ ID NO 50  
<211> LENGTH: 30  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 50

tccccccggg aaaagttaagt caaaaaggcac

30

<210> SEQ ID NO 51  
<211> LENGTH: 30  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 51

tccccccggg atggtctgct taaaatttcat

30

<210> SEQ ID NO 52  
<211> LENGTH: 45  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 52

tccccccggg tagcttgtac ccattaaaag aattttatca tgccg

45

-continued

<210> SEQ ID NO 53  
<211> LENGTH: 30  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 53

tccccccggg tttctcattc aagtggtaac

30

<210> SEQ ID NO 54  
<211> LENGTH: 30  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 54

tccccccggg taaataaaga aaataaaagtt

30

<210> SEQ ID NO 55  
<211> LENGTH: 47  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 55

aatttttgaa aattcaatat aaatggcttc agaaaaagaa attagga

47

<210> SEQ ID NO 56  
<211> LENGTH: 47  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 56

tcctaatttc ttttctgaa gccattata ttgaattttc aaaaatt

47

<210> SEQ ID NO 57  
<211> LENGTH: 51  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 57

agttttcacc aattggctcg cagccattat agtttttct ctttgacgtt a

51

<210> SEQ ID NO 58  
<211> LENGTH: 51  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 58

taacgtcaag gagaaaaaac tataatggct gcagaccaat tggtgaaaac t

51

-continued

<210> SEQ ID NO 59  
<211> LENGTH: 47  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 59

aattttgaa aattcaatat aaatgaaact ctcaactaaa ctttgtt 47

<210> SEQ ID NO 60  
<211> LENGTH: 47  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 60

aacaaagttt agttgagagt ttcatttata ttgaattttc aaaaattt 47

<210> SEQ ID NO 61  
<211> LENGTH: 47  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 61

aattttgaa aattcaatat aaatgtctca gaacgtttac attgttat 47

<210> SEQ ID NO 62  
<211> LENGTH: 47  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 62

ataacaatgtt aacggttctga gacatttata ttgaattttc aaaaattt 47

<210> SEQ ID NO 63  
<211> LENGTH: 51  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 63

tgcagaagtt aagaacggta atgacattat agtttttctt ctttgacgtt a 51

<210> SEQ ID NO 64  
<211> LENGTH: 51  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 64

taacgtcaag gagaaaaaac tataatgtca ttaccgttct taacctctgc a 51

&lt;210&gt; SEQ ID NO 65

-continued

<211> LENGTH: 47  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 65

aattttgaa aattcaatat aaatgtcaga gttgagagcc ttcagtg

47

<210> SEQ ID NO 66  
<211> LENGTH: 47  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 66

cactgaaggc tctcaactct gacatttata ttgaattttc aaaaatt

47

<210> SEQ ID NO 67  
<211> LENGTH: 51  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 67

ggtaacggat gctgtgtaaa cggtcattat agtttttct ctttgacgtt a

51

<210> SEQ ID NO 68  
<211> LENGTH: 51  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 68

taacgtcaag gagaaaaaac tataatgacc gtttacacag catccgttac c

51

<210> SEQ ID NO 69  
<211> LENGTH: 47  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 69

aattttgaa aattcaatat aaatgactgc cgacaacaat agtatgc

47

<210> SEQ ID NO 70  
<211> LENGTH: 47  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 70

gcatactatt gttgtcgca gtcatttata ttgaattttc aaaaatt

47

<210> SEQ ID NO 71  
<211> LENGTH: 70

-continued

<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 71

|   |    |
|---|----|
| gtaagacgg ttgggttta tctttgcag ttggtaactat taagaacaat cacaggaaac | 60 |
| agctatgacc  | 70 |

<210> SEQ ID NO 72  
<211> LENGTH: 70  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 72

|   |    |
|---|----|
| ttcggtttt tactttggtt cgctcaattt tgcaaggtaga taatcgaaaa gttgtaaaac | 60 |
| gacggccagt  | 70 |

<210> SEQ ID NO 73  
<211> LENGTH: 20  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 73

|                       |    |
|-----------------------|----|
| tgaagagcga gctcccgctg | 20 |
|-----------------------|----|

<210> SEQ ID NO 74  
<211> LENGTH: 24  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 74

|                            |    |
|----------------------------|----|
| tgaagagcct cgagttccct ttag | 24 |
|----------------------------|----|

<210> SEQ ID NO 75  
<211> LENGTH: 63  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 75

|   |    |
|---|----|
| ataatactagg aaaactcttt attaataac aatgatatat atatattcca gtggtgcatg | 60 |
| aac   | 63 |

<210> SEQ ID NO 76  
<211> LENGTH: 63  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 76

-continued

aattttgtgag tttagtatac atgcattac ttataataca gttttcaac gcattacgta 60

ggc 63

<210> SEQ ID NO 77  
<211> LENGTH: 24  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 77

cgctcgtcca acgcggcgcc acct 24

<210> SEQ ID NO 78  
<211> LENGTH: 24  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 78

acggccggcc aagcacgcgg ggat 24

<210> SEQ ID NO 79  
<211> LENGTH: 30  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 79

atggctacca atttgttggtg cttgtctaac 30

<210> SEQ ID NO 80  
<211> LENGTH: 36  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 80

tcagacatac atcaactgat taataggaaa agggtc 36

<210> SEQ ID NO 81  
<211> LENGTH: 24  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 81

atggctaccg agttgttggtg ctgt 24

<210> SEQ ID NO 82  
<211> LENGTH: 36  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 82

-continued

tcacctctca aaaggtaata tagttcagt aataac 36

<210> SEQ ID NO 83  
<211> LENGTH: 21  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 83

atggccaccc agttgttgtc c 21

<210> SEQ ID NO 84  
<211> LENGTH: 27  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 84

tcacctctca aaaggtaaga ttgggttc 27

<210> SEQ ID NO 85  
<211> LENGTH: 32  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 85

ttaaaaggat ccatggctac caatttggtg tg 32

<210> SEQ ID NO 86  
<211> LENGTH: 38  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 86

tatttcgcta gctcagacat acatcaactg attaatag 38

<210> SEQ ID NO 87  
<211> LENGTH: 32  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 87

ttaaaaggat ccatggctac cgagttggtg tg 32

<210> SEQ ID NO 88  
<211> LENGTH: 39  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 88

-continued

catatcgcta gtcacacctct caaaaaggtaa tataggttc

39

<210> SEQ ID NO 89  
<211> LENGTH: 30  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 89

aatttaggat ccatggccac cgagttgttgc

30

<210> SEQ ID NO 90  
<211> LENGTH: 39  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 90

tatatcgcta gtcacacctct caaaaaggtaa gattgggttc

39

<210> SEQ ID NO 91  
<211> LENGTH: 36  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 91

caaataaggat ccatggcaaa ttaccagcct aatttg

36

<210> SEQ ID NO 92  
<211> LENGTH: 38  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 92

caaataaggat ccatggctaa ttacgagcca aactcatg

38

<210> SEQ ID NO 93  
<211> LENGTH: 35  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 93

caaataaggat ccatggccaa ctacgagcca aattc

35

<210> SEQ ID NO 94  
<211> LENGTH: 42  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 94

gcgtacaaag gtaacttacc gtgataacc aaggcagcga cc

42

-continued

<210> SEQ ID NO 95  
<211> LENGTH: 47  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 95

ggctccccct ggccgcgtgt gagcttatat tgaatttca aaaattc 47

<210> SEQ ID NO 96  
<211> LENGTH: 44  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 96

gctcacacgc ggccaggggg agccatggct accgagttgt tgtg 44

<210> SEQ ID NO 97  
<211> LENGTH: 50  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 97

gctcacacgc ggccaggggg agccatggct aattacgagc caaactcatg 50

<210> SEQ ID NO 98  
<211> LENGTH: 48  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 98

aggtcggccg gcgttgacg agcgtcacct ctc当地 aatatagg 48

<210> SEQ ID NO 99  
<211> LENGTH: 44  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 99

ctagcgattc tcatttggaaat acgtcgacac tagtaataaca catc 44

<210> SEQ ID NO 100  
<211> LENGTH: 95  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 100

cgatgcgttc aattcctcta ctaatttagg gaaaacgttc aagaatctct ctctcataat 60

-continued

ttctttttctt gaagccatct ttgtatagcc cttaa

<210> SEQ ID NO 101  
<211> LENGTH: 44  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 101

gatgtgtatt actagtgtcg acgtattcca atgagaatcg ctag 44

<210> SEQ ID NO 102  
<211> LENGTH: 3901  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 102

gtttaaactt caaaagtcga tgcctcataa acttcggtag ttatattact ctgagatgac 60  
ttatactctt tttccaaatc cacattattt ggcgcaaagg tctcattgga agattccata 120  
agttggcagag agttcaatct ttttgaagag ccgcattaaat gtaatgatag attgtctggc 180  
attattccctt cctattctta ttatgcgtag gaatgtcttc gaaccgaaag atcttctcta 240  
tggggatgc ttttagtga aattaagaaa ggagttttat acagatgata cctaattatc 300  
atataagtaa gagagaacag agatthaatg gaaaatggaa aaggcataat tggcgctgaa 360  
tcaaataatgtt tattatatct ttacaatttg tcctgatttt gtccttgtct aacttggaaa 420  
tttttcatttc tgatgtcata cgactttttt ccggcttagg aaatcggtga aagctttttt 480  
ttttccatat ttcttgcctt atcggaaattt ttctgtcatt tctttcttc ctcgcgcgtt 540  
tctactaaaaa tctgaatttgt ccaaattcag tacaaaatta atcagtagga caaagggttc 600  
tgcgtatgcgtt cccggaaaaaaa aaaaggacaa aaagtttca acacggcaat ctctttttac 660  
acgtcaactgcgtt ccccacgtac aataacaaaaa tggtagtcat atcatgtcaa gaataggat 720  
ccaaaacgcga gcgggtgaaa gcatatcaag aattttgtcc ctgtgtttt aagtttgtgg 780  
ataatcgaaa tctcttacat tggaaacattt atcatataat catttattaa gtatgttgc 840  
catgtatgaa ctataaaaatgtt gttactactc gttatttttgcgtt gatgttgc 900  
ttatgtatgtt cgagaagtta agattatatg aataactaaa tactaaatag aatgttgc 960  
acagttagaaa caaaaacaaaaaaa aaaaacgcac agagaaacta aatccacatt aatttgagat 1020  
tctatctattt agaaaatgcac aactccactt aatggaaa acagataacc tctttttttt 1080  
ttttttatgtt tttgtatattc gaggctttt cttttgttag gtttatatttca atcatttca 1140  
tgaataaaaatgtt aagcttctta ttttgggtgc aaagaatgaa aaaaaggat ttttcatac 1200  
ttctaaatgtt tcaatttataa cccaaaattttataaaatgtt gggaaaatctgtatgtt 1260  
caagttaaac ctattccattt gcccctcgac gagtgctggg ggggtcggtttt ccactatcg 1320  
cgagttacttc tacacagccatc tgggtccaga cggccgcgtt tctggggcgat ttttgttac 1380  
ggccgcacatgtt cccggatcccg gatcgacgatc ttgcgtcgatc tgcacccctgc gccccaaatgtt 1440  
catcatcgaa attggcgatc accaagctt gatagagtttgcgtt gtcacagatc atgcggatc 1500  
tatacgcggccggc gagccgcggc gatctgtt gctccggatc cttccgcgtt gggaaatgtt 1560  
tatacgcggccggc gagccgcggc gatctgtt gctccggatc cttccgcgtt gggaaatgtt 1620

-continued

tctgctgctc cataacaagcc aaccacggcc tccagaagaa gatgttggcg acctctgtatt  
gggaatcccc gaacatcgcc tcgetccagt caatgaccgc tgtttatgcgg ccattgtccg  
tcaggacatt gttggagccg aaatcccggt gcacgagggt ccggacttcg gggcagtcct  
cgccccaaag catcagctca tcgagagccct gegrcaegga cgcactgacg gtgttgtcca  
tcacagttt ccagtgatac acatggggat cagcaatcgc gcataatgaaa tcacgcccatt  
tagtgttattt accgattctt tgccgtccga atggggccaa cccgctcg tcggctaaat  
cgcccgccagc gatecgatcc atggcctccg cgaccggctg cagaacagcg ggcagttccg  
tttcaggcag gtcttgcaac gtgacacccct gtgcacggcg ggagatgcaa taggtcaggc  
tctcgctgaa ttccccaaat tcaagcactt ccggaatcg gagcggccg gatgaaat  
gccgataaac ataacgatct ttgtagaaac catcgccgca gctattacc cgcaggacat  
atccacgccc tcctacatcg aagctgaaag cacgagattt tcgcctcc gagagctgca  
tcaggtcgga gacgctgtcg aactttcga tcagaaactt ctcgacagac gtcgcccgt  
gttcaggcattt tttcatttt aatgttactt ctcttgcaat taggaaacta taatgttaact  
caaataaga taaaacaaac taaaataaaa agaagttata cagaaaaacc cataaaacc  
agactaaatc cataataata atacacaaaaaa aactatcaa ataaaaccag aaaacagatt  
gaatagaaaaa atttttcga tctccatttta tattcaaatat tcgatataatg aaaaaggaa  
ctctcagaaa atcaccaaat caatttaattt agattttctt ttccttcta gcgttggaaa  
aaaaaattttt tctttttttt ttttagaaatg aaaaattttt gccgttagaa tcaccgtata  
aaccctgtat aaacgctact ctgttccactt gtgttaggcta tgattgaccc agtgttcatt  
gttattgcga gagagcggga gaaaagaacc gatacaagag atccatgctg gtatagttgt  
ctgtccaaaca ctttgatgaa cttgttaggac gatgtatgtt attactatgt tcgacactgc  
tgaagaattt gattttctt gccattccca tagacgttac aatccactaa ccgattcatg  
gatcttagtt tctccacaca gagctaaaag accttggta ggtcaacagg aggctgctta  
caagcccaca gctccattgt atgatccaaa atgctatcta tgcctggta acaaaaagac  
tactggtaac ctaaaccacaa gatataatc aacgtatattt ttccccaaatg attatgtgc  
cgtaggctc gatcaacactt ttttaccaca gaatgttcc aatgaggata atctttaaaa  
taggctgtt aaagtgcattt ctgtgagagg caatttttc gtcataatgtt tttagccccaa  
tcataatcta accattccac aatgaaaca atcagatctg gtcataatgtt ttaattcttgc  
gcaagcattt actgacgatc tctccagaga agcaagagaa aatcataagc cttccaaata  
tgtccaaata tttgaaaaca aaggtacagc catgggtgt tccaaacttac atccacatgg  
ccaagcttgg tgcttagaat ccattccatg tgaagtttcg caagaattga aatctttgt  
taaatataaaa cgtgaacacaa atactgtt gtttgcgcgt tacgtcaaat tagaatcaag  
agagaagtca agagtcgtag tggagaatga atccattttt gttgttggc cataactggc  
catctggcca tttgagaccc tggtcatttc aaagaagaag ctgcctcaa ttagocaatt  
taaccaaatg gtgaaggagg acctcgccctc gatTTTaaag caactaacta ttaagtatgt  
taatTTTTTaaat gaaacgagtt tccctactc aatgggtatc catcaggctc ctttgaatgc  
gactgggtat gaaatggatgat atagttgtt tcacatgtat ttctaccac ctttactgt  
atcagctact gttcgaaat tcttggttgg ttttgaattt gtaggtgagc ctgcattttaaa  
c 3901

-continued

<210> SEQ ID NO 103  
<211> LENGTH: 1118  
<212> TYPE: DNA  
<213> ORGANISM: *Saccharomyces cerevisiae*

<400> SEQUENCE: 103

|  |      |
|--|------|
| accataaccac agctttcaa ttcaattcat cattttttt ttattcttt ttttgattc     | 60   |
| ggtttcttg aaatttttt gattcgtaa tctccgaaca gaaggaagaa cgaaggaagg     | 120  |
| agcacagact tagattggtatataacgca tatgttagtgt tgaagaaaca tgaaattgcc   | 180  |
| cagtattctt aacccaactg cacagaacaa aaacctgcag gaaacgaaga taaatcatgt  | 240  |
| cgaagactac atataaggaa cgtgctgcta ctcatcttag tcctgttgct gccaagctat  | 300  |
| ttaatatcat gcacgaaaag caaacaaact tgtgtgtttc attggatgtt cgtagcc     | 360  |
| aggaattact ggagtttagtt gaagcattag gtcccaaaat ttgtttacta aaaacacatg | 420  |
| tggatatctt gactgatttt tccatggagg gcacagttaa gcccctaaag gcattatccg  | 480  |
| ccaagtacaa tttttactc ttcaagaca gaaaattgc tgacatttgt aatacagtca     | 540  |
| aattgcagta ctctgggggt gtatacagaa tagcagaatg ggcagacatt acgaatgcac  | 600  |
| acgggtgtggt gggcccaggt attgttagcg gttgaagca ggcggcagaa gaagtaacaa  | 660  |
| aggaacctag aggccctttt atgttagcag aattgtcatg caagggctcc ctatctactg  | 720  |
| gagaatatac taagggtact gttgacattg cgaagagcga caaagattt gttatcgct    | 780  |
| ttattgctca aagagacatg ggtggaaagag atgaaggtaa cgattgggtt attatgacac | 840  |
| ccgggtgtggg tttagatgac aaggagacg cattgggtca acagtagata accgtggatg  | 900  |
| atgtggtctc tacaggatct gacattatta ttgttggaaag aggactatggaa gcaaggaa | 960  |
| gggatgtcaa ggttaggggt gaacgttaca gaaaaggcgg ctgggaagca tatttgagaa  | 1020 |
| gatgcggcca gcaaaactaa aaaactgtat tataagtaaa tgcatgtata cttaactcac  | 1080 |
| aaatttagagc ttcaatttaa ttatatacgat tattaccc                        | 1118 |

<210> SEQ ID NO 104  
<211> LENGTH: 70  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 104

|   |    |
|---|----|
| tccttattac tgcgtatatac agtgtgaggt attctaagcg gttatattcac ctgcagaagt | 60 |
| taagattata  | 70 |

<210> SEQ ID NO 105  
<211> LENGTH: 46  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 105

|  |    |
|--|----|
| atccccgcgt gcttggccgg ccgtttcgac actagtaata cacatc | 46 |
|--|----|

<210> SEQ ID NO 106  
<211> LENGTH: 45  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence

-continued

<220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 106

atcggegctc gcggcctgca ggtagagct cctcgagaag ttaag 45

<210> SEQ ID NO 107

<211> LENGTH: 48

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 107

aacctgcagg ccgcgagcgc cgatatataat gtgtactttt cagttatg 48

<210> SEQ ID NO 108

<211> LENGTH: 46

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 108

cgaggaactc ttggatttct tgccacgact catctccatg cagttt 46

<210> SEQ ID NO 109

<211> LENGTH: 63

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 109

ataatactagg aaaactcttt atttaataac aatgatataat atatattcca gtggtgcatg 60

aac 63

<210> SEQ ID NO 110

<211> LENGTH: 63

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 110

aatttgtgag tttagtatac atgcatttac ttataataca gttttcaac gcattacgtt 60

ggc 63

<210> SEQ ID NO 111

<211> LENGTH: 65

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 111

ttctgtaatg aaggagaaaa ctcaccgagg cagttccata ggatgaccat accacagttt 60

ttcaa 65

-continued

<210> SEQ ID NO 112  
<211> LENGTH: 65  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 112

gccgcgcccc tgtagagaaa tataaaaggt taggattgc cactgggta ataaactgata 60  
taatt 65

<210> SEQ ID NO 113  
<211> LENGTH: 20  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 113

tccgactctag gcagataagg 20

<210> SEQ ID NO 114  
<211> LENGTH: 20  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 114

gtcgataacc atggtaactcc 20

<210> SEQ ID NO 115  
<211> LENGTH: 20  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 115

atccctaagtc gtggctattg 20

<210> SEQ ID NO 116  
<211> LENGTH: 20  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 116

accgaatgat gctctagatg 20

-continued

<210> SEQ ID NO 117  
<211> LENGTH: 70  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 117

|   |    |
|---|----|
| tggcttcaga aaaagaaatt aggagagaga gattcttcaa | 60 |
| agcttttcaa                                  | 70 |

<210> SEQ ID NO 118  
<211> LENGTH: 70  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 118

|   |    |
|---|----|
| tatTTTCAA taaagaatat ctccactac tgccatctgg cgtcataact gggtaataac | 60 |
| tgatataatt  | 70 |

<210> SEQ ID NO 119  
<211> LENGTH: 32  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 119

|                                     |    |
|-------------------------------------|----|
| attactgcag gacggtagca acaagaatat ag | 32 |
|-------------------------------------|----|

<210> SEQ ID NO 120  
<211> LENGTH: 41  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 120

|   |    |
|---|----|
| gaaataggca aagttgacat ttttgaggga atattcaact g | 41 |
|---|----|

<210> SEQ ID NO 121  
<211> LENGTH: 41  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 121

|   |    |
|---|----|
| cagttaata ttccctcaaa aatgtcaact ttgccttattt c | 41 |
|---|----|

<210> SEQ ID NO 122  
<211> LENGTH: 32  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 122

-continued

|                                     |    |
|-------------------------------------|----|
| tataggatcc ctgcgagcgt cccaaaacct tc | 32 |
|-------------------------------------|----|

```

<210> SEQ ID NO 123
<211> LENGTH: 44
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      primer

```

<400> SEQUENCE: 123

|  |    |
|--|----|
| ttgtgatgct aaagttatga gtctcgagaa gttaagatta tatg | 44 |
|--|----|

```

<210> SEQ ID NO 124
<211> LENGTH: 44
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      primer

```

<400> SEQUENCE: 124

|  |    |
|--|----|
| catataatct taacttctcg agactcataa ctttagcatc acaa | 44 |
|--|----|

```

<210> SEQ ID NO 125
<211> LENGTH: 28
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      primer

```

<400> SEQUENCE: 125

|                                  |    |
|----------------------------------|----|
| gtttaaacatt caaagctcg a tgcctcat | 28 |
|----------------------------------|----|

```

<210> SEQ ID NO 126
<211> LENGTH: 28
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      primer

```

<400> SEQUENCE: 126

|                                 |    |
|---------------------------------|----|
| gtttaaacgca ggctcaccta acaattca | 28 |
|---------------------------------|----|

```

<210> SEQ ID NO 127
<211> LENGTH: 44
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      primer

```

<400> SEQUENCE: 127

|  |    |
|--|----|
| gatgtgtatt actagtgtcg acactgctga agaatttgat tttt | 44 |
|--|----|

```

<210> SEQ ID NO 128
<211> LENGTH: 44
<212> TYPE: DNA

```

-continued

<213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

&lt;400&gt; SEQUENCE: 128

aaaaatcaa ttcttcagca gtgtcgacac tagtaataca catc

What is claimed is:

1. A gaseous isoprene composition comprising isoprene, carbon dioxide and water, wherein the isoprene is in an amount greater than about 10% by volume; wherein the carbon dioxide is in an amount greater than about 10% by volume; wherein the water is in an amount greater than about 70% of its saturation amount; and wherein the gaseous isoprene composition comprises 1 part per million or less of C<sub>2</sub>-C<sub>5</sub> alkynes.

2. The gaseous isoprene composition of claim 1, wherein the gaseous isoprene composition further comprises less than about 15% oxygen by volume.

3. The gaseous isoprene composition of claim 1, wherein the gaseous isoprene composition further comprises less than about 10% oxygen by volume.

4. The gaseous isoprene composition of claim 1, wherein the gaseous isoprene composition further comprises less than about 5% oxygen by volume.

5. The gaseous isoprene composition of claim 1, wherein the gaseous isoprene composition further comprises less than about 1% oxygen by volume.

6. The gaseous isoprene composition of claim 1 comprising less than 3% water by weight.

7. The gaseous isoprene composition of claim 1 comprising greater than about 20% carbon dioxide by volume.

8. The gaseous isoprene composition of claim 1 comprising greater than about 30% carbon dioxide by volume.

9. The gaseous isoprene composition of claim 1, wherein the gaseous isoprene composition further comprises nitrogen in an amount greater than about 10% by volume.

10. The gaseous isoprene composition of claim 1, wherein the gaseous isoprene composition further comprises argon in an amount less than about 0.9% by volume.

15 11. The gaseous isoprene composition of claim 1, wherein the gaseous isoprene composition further comprises ethanol in an amount less than about 0.5% by volume.

12. The gaseous isoprene composition of claim 1, wherein the gaseous isoprene composition further comprises ethanol in an amount more than 1% by volume.

20 13. The gaseous isoprene composition of claim 1, wherein the isoprene is in an amount greater than about 10% by volume; wherein the carbon dioxide is in an amount greater than about 30% by volume; wherein the water is in an amount greater than about 70% of its saturation amount; wherein oxygen is in an amount less than about 5% by volume; and wherein the gaseous isoprene composition comprises 1 part per million or less of C<sub>2</sub>-C<sub>5</sub> alkynes.

30 14. The gaseous isoprene composition of claim 1, wherein the gaseous composition further comprises 1 part per million or less of cyclopentadiene.

15 15. The gaseous isoprene composition of claim 1, wherein the gaseous composition further comprises 1 part per million or less of piperylene.

35 16. The gaseous isoprene composition of claim 1, wherein the gaseous composition further comprises 1 part per million or less of 1,4-pentadiene.

\* \* \* \* \*

UNITED STATES PATENT AND TRADEMARK OFFICE  
**CERTIFICATE OF CORRECTION**

PATENT NO. : 9,233,894 B2  
APPLICATION NO. : 13/887381  
DATED : January 12, 2016  
INVENTOR(S) : Derek McPhee

Page 1 of 1

It is certified that error appears in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Claims

In claim 1, column 139, line 19, please change “C2-05 alkynes” to --C2-C5 alkynes--.

Signed and Sealed this  
Nineteenth Day of April, 2016



Michelle K. Lee  
*Director of the United States Patent and Trademark Office*